WAHRNEHMUNG UND REPRÄSENTATION MUSIKALISCHER STRUKTUREN

Funktionale und strukturelle Aspekte eines kognitiven Modells des Musikhörens

INAUGURAL-DISSERTATION

zur

Erlangung des Grades eines Doktors der Philosophie

in der

Abteilung für Philosophie, Pädagogik, Psychologie

der

RUHR-UNIVERSITÄT BOCHUM

vorgelegt von

THOMAS H. STOFFER

Bochum 1981

WAHRNEHMUNG UND REPRÄSENTATION MUSIKALISCHER STRUKTUREN

Funktionale und strukturelle Aspekte eines kognitiven Modells des Musikhörens

INAUGURAL-DISSERTATION

zur

Erlangung des Grades eines Doktors der Philosophie

in der

Abteilung für Philosophie, Pädagogik, Psychologie

der

RUHR-UNIVERSITÄT BOCHUM

vorgelegt von

THOMAS H. STOFFER

Bochum 1981

Gedruckt mit Genehmigung der Fakultät der Abteilung für
Philosophie, Pädagogik, Psychologie der Ruhr-Universität Bochum

Referent: Priv. Doz. Dr. Thomas Schulz
Korreferent: Prof. Dr. Antonius van der Geest

Tage der mündlichen Prüfungen: 27. Mai und 3. Juni 1981

UNIV.
LEIDEN
BBL

VORWORT

Allen, die mich auf dem manchmal beschwerlichen Weg bis zur Fertigstellung dieser Untersuchung begleitet haben, sei es durch tatkräftige Unterstützung oder "nur" durch ihren Zuspruch oder ihre kritische Stellungnahme, sei an dieser Stelle gedankt:

Zuallererst möchte ich Herrn Prof. Dr. O. Graefe danken, der mir die Möglichkeit bot, an seinem Forschungsprojekt mitzuarbeiten, und damit die materiellen Voraussetzungen schuf, ohne die diese Untersuchung nicht zustande gekommen wäre. Mein Dank gilt in gleicher Weise dem Minister für Wissenschaft und Forschung des Landes Nordrhein-Westfalen, der durch die Förderung des Forschungsprojekts "Experimentalpsychologische Analyse kognitiver Prozesse beim Hören, Lernen und Behalten von Musik in Abhängigkeit von der syntaktischen Struktur der Musik", zu dem auch unsere Untersuchung unter vielen anderen einen Beitrag leisten konnte, die finanziellen Mittel für kostspielige Apparaturen und die Bezahlung eines Teils der Versuchspersonen bereitstellte.

Unter denen, die mich bei der zeitraubenden Arbeit des Experimentierens und Auswertens von Daten unterstützt haben, seien hier insbesondere Frau Petra Flaßhofe, Frau Rohmeder-Schmitz und Herr Rudolf Vogelhuber hervorgehoben. Herrn Vogelhuber bin ich zu besonderem Dank verpflichtet, weil er durch seine verantwortungsbewußte und selbständige Mitarbeit als Versuchsleiter, Organisator und als kompetenter Musiker wesentlich zum Erfolg der Experimente beigetragen hat.

Und nicht zuletzt möchte ich meiner Frau für den Zuspruch, ihre Anregungen und die Geduld danken, mit der sie die Entwicklung dieser Arbeit begleitet hat, und für die Anfertigung der Abbildungen, die sie trotz ihrer knapp bemessenen Freizeit für mich zeichnete.

Ein Teil der Befunde wurde unter dem Titel "Zur 'psychologischen Realität' einer musikalischen Syntax" auf dem 1. Bochumer kognitionspsychologischen Kolloquium (31.1o. - 2.11.198o in Altena) vorgetragen. Eine vorläufige Fassung der in Teil II der Untersuchung verwendeten Syntax deutscher Volks- und Kinderlieder ist als Bericht 11/1979, Psychologisches Institut der Ruhr-Universität Bochum, Arbeitseinheit Kognitionspsychologie, erschienen.

Kleve, im März 1981

Thomas H. Stoffer

INHALTSVERZEICHNIS

ALLGEMEINE EINLEITUNG 1

Strukturelles Hören von Musik (4) - Analogien zwischen Sprache und Musik (6) - Zur "psychologischen Realität" musikalisch-syntaktischer Strukturen (9) - Plan der Untersuchung (12)

TEIL I

AUDITIVE MUSTERERKENNUNG: WAHRNEHMUNG UND REPRÄSENTATION TONALER STRUKTUREN 16

1. EINLEITUNG UND ÜBERBLICK 17

2. AUDITIVE INFORMATIONSVERARBEITUNG DER STRUKTUR VON TONFOLGEN 22

2.1 AUDITIVE INFORMATIONSVERARBEITUNG: CHARAKTERISIERUNG DES ANSATZES 24

2.2 MUSTERERKENNUNGSLEISTUNGEN BEIM HÖREN VON TONFOLGEN UND DIE IHNEN ZUGRUNDELIEGENDEN PROZESSE DER INFORMATIONSVERARBEITUNG 28

2.2.1 ERKENNEN DER BEZIEHUNGEN ZWISCHEN TÖNEN 29

Erkennen der zeitlichen Reihenfolge in Abhängigkeit vom SOA (3o) - Erkennen tonaler Muster in Abhängigkeit von der Größe und Konfiguration der Intervalle (42) - Elementare Merkmale (45) - Konfigurative Merkmale (46)

2.2.2 ERKENNEN DER MELODIEKONTUR 47

2.2.3 ORGANISATION BEI WAHRNEHMUNG UND SPEICHERUNG VON TONFOLGEN 52

Integration und Segmentierung (53) - Bedingungen für die Integration sukzessiver Wahrnehmungen (56) - Zeitliche Nähe und Kohärenz (56) - Rhythmus und prosodische Struktur (58) - Ähnlichkeiten zwischen Tönen aufgrund ihrer Tonhöhe (63) - Wissensstrukturen und ihre Funktionen bei der Integration und Segmentierung (69)

2.3 ZUSAMMENFASSUNG UND FOLGERUNGEN 74

3. STRUKTURELLE UND FUNKTIONALE ASPEKTE DER REPRÄSENTATION KOMPLEXER REIZMUSTER 79

3.1 GESTALTTHEORIE UND MUSIKWAHRNEHMUNG 81

3.2 VERARBEITUNG VON STRUKTUREN UND KONFIGURATIONEN 87

Effekte der Konfiguration bei der visuellen Identifikation (88) - Globale und lokale Reizmusteranalyse (92)

3.3 STRUKTURELLE UND FUNKTIONALE ASPEKTE DER KOGNITIVEN REPRÄSENTATION 97

Das Hierarchiekonzept (97) - Das hierarchische Repräsentationsebenenmodell von Wolff (1977) (99)

3.4 ZUSAMMENFASSUNG 113

4. AUDITIVES CHUNKING: DIE BILDUNG FUNKTIONALER EINHEITEN DER AUDITIVEN ANALYSE 115

4.1 BILDUNG VON REPRÄSENTANZEN UND AKTIVIERUNG STRUKTURELLER REPRÄSENTATIONSNIVEAUS BEI DER AUDITIVEN WAHRNEHMUNG 118

Spezifizierung des Hierarchiemodells für die Wahrnehmung auditiver Reizmuster (119) - Funktionale Bedeutung supraprimärer Repräsentationsniveaus beim Musikhören (125)

4.2 VERARBEITUNGSEINHEITEN INNERHALB DER PSYCHISCHEN PRÄSENZZEIT 13o

4.3 AUDITIVES CHUNKING 136

Empirische Belege für ein auditives Chunking (137) - Erklärung differentieller Schwierigkeiten bei der auditiven Mustererkennung durch das Modell des auditiven Chunkings und das der Repräsentationshierarchie (14o)

4.4 ZUSAMMENFASSUNG 148

5. EXPERIMENT 1: FOKUSSIERUNG STRUKTURELLER DETAILS EINER TONFOLGE IN ABHÄNGIGKEIT VOM VORAUSGEGANGENEN TONALEN KONTEXT 152

5.1 EINLEITUNG 152

5.2 METHODE 158

Reizmaterial (158) - Versuchsapparaturen (158) - Versuchsplan (159) - Versuchsdurchführung (16o) - Instruktion (161) - Versuchspersonen (162)

5.3 BEFUNDE UND DISKUSSION 162

Auswertungsverfahren (162) - Vergleich der experimentellen und Kontrollbedingungen (163) - Der Einfluß der Gesamtkonfiguration auf die Fokussierbarkeit des letzten Intervalls (165) - Der Einfluß der zeitlichen Gruppierung der Töne auf die Fokussierbarkeit des letzten Intervalls (178) - Abhängigkeit der Konturdiskrimination vom SOA zwischen den beiden letzten Tönen (182)

5.4 ZUSAMMENFASSUNG 185

6. EXPERIMENT 2: WAHRNEHMUNG UND IDENTIFIKATION STRUKTURELLER INVARIANZEN IN TONFOLGEN 189

6.1 EINLEITUNG 189

Das Invarianzkonzept Gibson's (192) - Repräsentation von Invarianzstrukturen (194) - "Tuning" als Fourieranalyse (197) - Identifikation struktureller Invarianzen (2o4)

6.2 METHODE 2o9

Versuchsplan (2o9) - Reizdarbietung und Aufgabenstellung (21o) - Versuchsablauf (212) - Versuchspersonen (214)

6.3 BEFUNDE UND DISKUSSION 214

Auswertungsverfahren (214) - Haupteffekt der Intervallkonfiguration (215) - Effekt des Faktors "Gesamt-vs. Teilaufgabe" (217) - Der Effekt der Faktoren SOA und Δf (22o) - Unterschiede zwischen "same"- und "different"-Bedingungen (226)

6.4 ZUSAMMENFASSUNG 232

7. ALLGEMEINE DISKUSSION 235

Struktur und Format der Repräsentation von Tonfolgen und von Musik (235) - Entdecken von Invarianzen und Fourieranalyse (244) - Strukturelle Identifikation durch Resonanz (246)

TEIL II

WAHRNEHMUNG UND REPRÄSENTATION MUSIKALISCHER STRUKTUREN: ZUR PSYCHOLOGISCHEN BRAUCHBARKEIT EINER MUSIKALISCHEN SYNTAX

1. EINLEITUNG 253

1.1 KOGNITIVE MUSIKPSYCHOLOGIE: EIN FORSCHUNGSPROGRAMM 254

1.2 METHODEN DER FORMALEN STRUKTURBESCHREIBUNG VON MUSIK 261

1.2.1 INFORMATIONSTHEORIE DER MUSIK 263

1.2.2 KODIERSYSTEME FÜR SERIELLE MUSTER 267

1.2.3 ÜBERTRAGUNG LINGUISTISCHER METHODEN AUF DIE ANALYSE UND BESCHREIBUNG MUSIKALISCHER STRUKTUREN 27o

Analogien zwischen Sprache und Musik (271) - Musiktheoretische Anknüpfungspunkte (274) - Erste Ansätze zur Formulierung musikalisch-syntaktischer Regeln (277)

1.3 DESKRIPTIVE UND KOGNITIVE ADÄQUATHEIT DER SYNTAX DEUTSCHER VOLKS- UND KINDERLIEDER 282

Beschreibung der Syntax (283) - Deskriptive Adäquatheit der Syntax (286) - Kognitive Adäquatheit der Syntax (288)

1.4 ZUSAMMENFASSUNG 293

2. EXPERIMENTE ZUR REPRÄSENTATIONSSTRUKTUR BEIM HÖREN VON MELODIEN 295

2.1 SYNTAXTRAINING 296

Beschreibung des Trainingsprogramms (298) - Durchführung des Trainings (299)

2.2 EXPERIMENT 3: LOKALISATION EINES EINER MELODIE ÜBERLAGERTEN KLICKS 3oo

2.2.1 EINLEITUNG 3oo

Die Anwendung des Klickparadigmas in der psycholinguistischen Forschung (3o2) - Übertragung des Klickparadigmas auf die Untersuchung der Repräsentationsstruktur beim Hören von Melodien (314) - Erklärung des Fehllokalisationseffekts im Rahmen des Hierarchiemodells (316)

2.2.2 METHODE 32o

Reizmaterial und Darbietungsweise (32o) - Versuchsplan (322) - Versuchsapparaturen (323) - Versuchsdurchführung (323) - Versuchspersonen (325)

2.2.3 BEFUNDE UND DISKUSSION 325

Auswertungsmethode (325) - Klicklokalisationshäufigkeiten (326) - Klickverschiebungsscores (329)

2.3 EXPERIMENT 4: VERGLEICH SYNTAKTISCH TRAINIERTER UND UNTRAINIERTER HÖRER IN EINER KLICKDETEKTIONSAUFGABE 333

2.3.1 EINLEITUNG 333

2.3.2 METHODE 338

Reizmaterial, Versuchsapparaturen und Darbietungsweise (338) - Versuchsplan und Versuchspersonen (339) - Versuchsdurchführung (34o) -

2.3.3 BEFUNDE UND DISKUSSION 34o

Auswertungsmethode (34o) - Haupteffekte (341) - Interaktion zwischen Vpn-Gruppen und Klickpositionen (344)

2.4 EXPERIMENT 5: SAME-DIFFERENT-RT BEI VERGLEICH SYNTAKTISCH DEFINIERTER MELODIEAUSSCHNITTE IN ABHÄNGIGKEIT VON IHRER SYNTAKTISCHEN UND ZEITLICHEN NÄHE 347

2.4.1 EINLEITUNG 347

2.4.2 METHODE 353

Reizmaterial und Darbietungsweise (353) - Versuchsplan (354) - Versuchsdurchführung (354) - Versuchspersonen (357)

2.4.3 BEFUNDE UND DISKUSSION 357

Auswertungsmethode (357) - Bedingungen syntaktischer und zeitlicher Nähe (357) - Bedingungen unterschiedlichen Vergleichsniveaus (359)- Effekt der same- und different-Bedingungen (362)

2.5 ZUSAMMENFASSUNG DER BEFUNDE AUS DEN EXPERIMENTEN 3 BIS 5 363

3. ALLGEMEINE DISKUSSION 367

Ernst Kurth's Musikpsychologie (367) - Kognitiv adäquate Aspekte der Syntax (368)- Funktionen musikalischen Wissens für die Wahrnehmung von Musik (374) - Versuch einer kognitionspsychologischen Beschreibung des strukturellen Hörens (387) - Kognitive Prozesse bei nicht primär auf die musikalische Struktur gerichteten Höreinstellungen (391) - Abschließende Bemerkung (394)

ZUSAMMENFASSUNG 396

LITERATURVERZEICHNIS 4o4

ANHANG 1: Syntax der in den Experimenten verwandten Melodien 433

ANHANG 2: Syntaxtraining 499

ANHANG 3: Varianzanalysen 561

ALLGEMEINE EINLEITUNG

Musik, aber auch andere Bereiche der Kunst, haben zeitweise eine Art übernatürlicher Qualität angenommen, die es als Sakrileg erscheinen lassen mußte, wenn man versuchen wollte, sich ihr mit dem analytischen Werkzeug einer rationalen Wissenschaft zu nähern. Kennzeichnend für eine solche "erhabene" Sicht von Musik ist die poetisierende Beschreibungsweise der musikalischen Hermeneutiker (Kretzschmar, 19o2; Schering, 1934). Der Argwohn gegenüber einer rationalen Auseinandersetzung mit dem Gegenstand Musik, der sich hinter diesem musikwissenschaftlichen Ansatz verbirgt, ist in der Ansicht begründet, Musik sei als "Sprache der Emotionen" (Pratt, 1952) zu begreifen, und jede rationale Beschäftigung mit Musik ginge an ihrem "Wesen" vorbei.

Dieses Dogma prägte lange Zeit auch die Musikpsychologie bis in die ausgehenden sechziger Jahre (z.B. Capurso et al., 1972; Hevner, 1935, 1936; Kleinen, 1968; Künding, 1961; Schoen, 1927). Das zentrale Interesse am musikalischen Erleben ist mit der weit verbreiteten Dominanz einer Höreinstellung verbunden, die Adorno (1968) dem Typus des e m o t i o n a l e n H ö r e r s zuschreibt und die in vielen anderen Typologien musikalischer Höreinstellungen ein besonderes Gewicht besitzt (z.B. der "Emotions- und Gefühlstypus" bei Alt (1935); der "Charaktertypus" bei Myers (1922; Myers & Valentine, 1914); das "pathetische Erleben" bei Nadel (1931)). Sie läßt sich durch die Bereitschaft charakterisieren, in der Musik den Ausdruck von Affekten zu sehen, von denen sich der Hörer gerne gefangennehmen läßt (zur Änderung der Stimmungslage durch Musik vgl. Shatin (197o)).

Diese Höreinstellung ist, wenn sie dominiert, nicht dazu geeignet, Musik adäquat wahrzunehmen (Adorno, 1968; Rauhe et al., 1975), weil in ihr mehr das eigene Erleben als die Musik Gegenstand der Wahrnehmung wird. Die Ausschließlichkeit, mit der diese Form des Musikhörens lange Zeit zum Gegenstand musikpsychologischer Forschung erhoben wurde, engte den wissenschaftlichen Horizont der Musikpsychologie so stark ein, daß der Kontakt zur allgemeinen Wahrnehmungspsychologie, zu der die Musikpsychologie als ein von ihrem Gegenstand her spezialisierter Zweig eigentlich zu rechnen ist, verloren zu gehen droht. Sie gerät dabei in Gefahr, die Subjektseite zu Ungunsten der Objektseite überzubetonen und damit Abhängigkeitsbeziehungen zwischen beiden so undifferenziert zu sehen, daß die Musik zum beliebig austauschbaren Auslöser subjektiven Erlebens degradiert wird. Wahrnehmungspsychologische Fragestellungen im engeren Sinne wurden überwiegend in den Bereich der Psychoakustik verwiesen, die sich aber lediglich mit der Psychophysik isolierter akustischer Signale beschäftigt (vgl. Carterette & Friedman, 1978; Tobias, 197o), deren strukturelle Komplexität bei weitem nicht an die der Musik heranreicht. Damit wird es verständlich, daß ein Psychoakustiker seinen mit dem Titel "Musical Perception" überschriebenen Artikel mit einer Bemerkung abschließen muß, die dem Sinne nach ausdrückt, daß im Bereich der Musikwahrnehmung vieles noch ungeklärt sei (Ward, 197o).

Im Bereich der Psychologie der auditiven Wahrnehmung klafft ganz offensichtlich eine Erkenntnislücke zwischen den peripheren Prozessen der sensorischen Verarbeitung einzelner akustischer Signale und den emotionalen Prozessen des musikalischen Erlebens. Vitz (1974) kritisiert deshalb mit Recht die beinahe völlige Vernachlässigung kognitiver Verarbeitungsprozesse durch die experimentelle Musikästhetik und Musikpsychologie: Es werde lediglich der Input und der Output durch

den Hörer betrachtet, aber alles weitere, was dazwischen an Verarbeitung geschehe, werde weder theoretisch noch experimentell thematisiert. Einen Beitrag zur Schließung dieser Lücke zu leisten und einige Grundlagen für ein psychologisches Modell des Musikhörens zu legen, ist die Absicht dieser Untersuchung.

Man könnte vielleicht versucht sein, in einer Betonung kognitiver Prozesse beim Musikhören eine ebensolche Einseitigkeit zu vermuten, wie wir sie in der Betonung des emotionalen Erlebens von Musik zu erkennen glauben. Rauhe et al. (1975) sehen nämlich im Gegensatz zu Adorno (1968) in einer "cogitativen Rezeption" (Rauhe, 1968), die in einer Analyse struktureller Details der Musik besteht, keineswegs eine adäquate Rezeptionsweise. Ein solcher Einwand wäre aber nur dann gerechtfertigt, wenn man zu den kognitiven Prozessen nur solche rechnen wollte, wie Adorno (1968) sie bei der Charakterisierung des Expertentypus beschreibt:

"Der Experte selbst wäre ... durch gänzlich adäquates Hören zu definieren. Er wäre der voll bewußte Hörer, dem tendentiell nichts entgeht und der zugleich in jedem Augenblick über das Gehörte Rechenschaft sich ablegt. ... Während er dem Verlauf auch verwickelter Musik spontan folgt, hört er das Aufeinanderfolgende: vergangene, gegenwärtige und zukünftige Augenblicke so zusammen, daß ein Sinnzusammenhang sich herauskristallisiert" (Adorno, 1968, S. 15).

In dieser Beschreibung Adornos handelt es sich immer um b e w u ß t e Repräsentationen. Auch Rauhes (1968) Begriff der "cogitativen Rezeption" meint immer kognitive Prozesse des Erkennens und bewußte Repräsentation des Erkannten. Man würde aber einen logischen und psychologischen Fehler begehen, wenn man in der Negation dieser Beschreibung im Zusammenhang einer Charakterisierung andersartiger Rezeptionsweisen davon ausgehen würde, daß dort kognitive Prozesse bedeutungslos wären.

Dieser Fehlschluß würde in einer Gleichsetzung des prozessualen und Repräsentationsaspekts der Verarbeitung bestehen: Wenn eine dem musikalischen Ausdruck zugeschriebene Emotion repräsentiert wird, muß es dennoch kognitive Prozesse der Verarbeitung zumindest einiger oberflächlicher Strukturaspekte der Musik geben, wenn die Beziehung zwischen der musikalischen Struktur und dem repräsentierten Eindruck keine zufällige sein soll, was sie ja auch nicht ist (Hevner, 1934, 1935, 1936, 1937; Kleinen, 1968; Meyer, 1956; Vitz, 1966).

Kognitive Prozesse der Musikwahrnehmung zu thematisieren, heißt also nicht notwendigerweise, die Wahrnehmung von Musik nur unter einer ganz bestimmten Höreinstellung zu untersuchen, sondern vielmehr Prozesse zu untersuchen, die einer bewußten Repräsentation gleich welchen Inhalts vorausgehen und sie begleiten: An jeder Form musikalischer Wahrnehmung ist, wie auch immer der sich einstellende Eindruck oder der erkannte Strukturaspekt phänomenal in Erscheinung tritt, eine kognitive Verarbeitung zumindest einiger struktureller Aspekte beteiligt. Ein weitergehendes Verständnis der vielfältigen Repräsentationsformen beim Hören von Musik, wenn es über ein Konstatieren bewußt repräsentierter Inhalte hinausgehen soll, erfordert als Grundlage eine kognitive Theorie der Verarbeitung musikalischer Strukturen.

Strukturelles Hören von Musik

Eine kognitive Theorie des Musikhörens zu formulieren, die die Entstehung verschiedener Formen der Repräsentation zu erklären vermag, ist das Ziel dieser Untersuchung. Den Anspruch aber, eine solche Theorie in all ihren Aspekten experimentell zu überprüfen, können wir nicht erfüllen; dies würde den Rahmen dieser Untersuchung sprengen.

Daher besteht die Notwendigkeit, eine Form der Repräsentation von Musik in den Mittelpunkt zu stellen, um paradigmatisch wichtige theoretische Punkte einer experimentellen Überprüfung zu unterziehen. Wir konzentrieren unsere theoretischen Ausführungen und experimentellen Untersuchungen deshalb zunächst auf das "strukturelle Hören", da bei Vorliegen einer Repräsentation struktureller Aspekte der in einer experimentellen Situation zu hörenden Musik eine eindeutige Beziehung zwischen der wahrgenommenen musikalischen Struktur und dem Verhalten des Hörers hergestellt werden kann. Zum Schluß werden wir aber den theoretischen Rahmen wieder weiter spannen und andere Höreinstellungen und Repräsentationsformen berücksichtigen.

Keil (1974a) differenziert zwischen vier Bedeutungen des Begriffs "strukturelles Hören" : 1. "Strukturelles Hören" meint in seiner ursprünglichen Bedeutung das adäquate Hören struktureller bzw. serieller Musik. Diese eingeschränkte Verwendung des Begriffs ist aber schon deshalb kaum sinnvoll, weil nicht nur serielle Musik eine "Struktur" besitzt (zum musikalischen Strukturbegriff vgl. Keil (1974b)), und auch deshalb wenig aussagekräftig, weil hier der wertende Begriff des "adäquaten Hörens" verwendet wird, der selbst wiederum einer Definition bedarf. 2. "Strukturelles Hören" meint das Erfassen kleinster musikalischer Einheiten der Melodie, des Rhythmus, der Harmonik, Dynamik, Klangfarbe, des Tempos und der Phrasierung. 3. "Strukturelles Hören" als Erfassen weitgespannter musikalischer Entwicklungen von Motiven, Themen und deren Verarbeitung in ihrer linearen Abfolge und vertikalen Schichtung bezieht sich auf musikalische Einheiten größerer zeitlicher Erstreckung. 4. Schließlich kann "strukturelles Hören" auch bedeuten, daß der Hörer mit der Intention an das Musikstück herangeht, die identifizierten Einheiten geringer und größerer zeitlicher Erstreckung im Zusammenhang des Ganzen wahrzunehmen.

Beim strukturellen Hören handelt es sich also um eine Höreinstellung, bei der die Aufmerksamkeit auf die Wahrnehmung struktureller Aspekte der Musik gerichtet ist, deren Korngröße vom einzelnen Ton und seinen Merkmalen bis hin zum ganzen Musikstück reicht, wobei die fokussierten strukturellen Merkmale bewußt kognitiv repräsentiert werden (zum Begriff der "Fokussierung" vgl. Neumann (1980) und Wolff (1977)).

Analogien zwischen Sprache und Musik

Der Erfolg unseres Versuchs, strukturelles Hören von Musik kognitionspsychologisch zu erklären, steht und fällt mit der Adäquatheit der musikalischen Strukturbeschreibung, die der Erzeugung und Variation der in Experimenten darzubietenden Musik zugrunde gelegt wird. Da das strukturelle Hören auf die kognitive Repräsentation musikalischer Strukturaspekte gerichtet ist, bedarf es eines Modells der musikalischen Strukturbeschreibung, das gleichzeitig ein hypothetisches Modell der strukturellen Repräsentation von Musik ist. Das Modell der Strukturbeschreibung muß also wenigstens eine Doppelfunktion erfüllen: Es muß 1. eine objektive, die Möglichkeiten der systematischen Strukturvariation erkennen lassende, formale Beschreibung der musikalischen Struktur auf allen Ebenen der vertikalen Schichtung liefern und 2. gleichzeitig so formuliert sein, daß es den Inhalt kognitiver Repräsentationen bei gegebener Aufgabe für den Hörer in Form einer Hypothese spezifiziert, die eine Vorhersage des Verhaltens des Hörers bei der Aufgabenlösung ermöglicht.

Form und Inhalt einer kognitiven Repräsentation sind aber abhängig vom Wissen des Wahrnehmenden über Struktur und Identität der in einem bestimmten Umweltausschnitt anzutreffenden Objekte. Bruner

(1957) spricht in diesem Zusammenhang von "perceptual readiness"; Neisser (1967, 1976) würde sagen, daß der Wahrnehmende über geeignete kognitive Schemata verfügen muß, um ein Objekt seiner Umwelt identifizieren und in seiner Funktion für die ablaufende Handlung repräsentieren zu können. Hieraus ergeben sich für unsere Untersuchung zwei Konsequenzen: 1. Die musikalische Strukturbeschreibung muß über die genannte Doppelfunktion hinaus auch noch die Funktion erfüllen, das beim Hörer potentiell vorhandene Wissen über musikalische Strukturen zu beschreiben. 2. Wir müssen unsere Untersuchung auf einen bestimmten musikalischen Formtyp einschränken, über dessen Strukturen unsere potentiellen Hörer bereits ein möglichst differenziertes Grundwissen besitzen. Aus diesem Grunde beschränken wir unsere experimentellen Untersuchungen auf die Wahrnehmung deutscher Volks- und Kinderliedmelodien, weil wir bei diesem musikalischen Formtypus davon ausgehen können, daß alle unsere Versuchspersonen durch ihre musikalischen Erfahrungen in der Kindheit in diesem Bereich über ein zumindest implizites Wissen (Polanyi, 1966) verfügen, das in einem Trainingsprogramm leicht zu einem differenzierten Wissen erweitert werden kann.

Den einzuschlagenden Weg zu einem Modell der musikalischen Strukturbeschreibung, das die drei genannten Anforderungen erfüllen könnte, weist eine These, die Adorno (1963) besonders prägnant formuliert:

"Musik ist sprachähnlich. Ausdrücke wie musikalisches Idiom, musikalischer Tonfall, sind keine Metaphern. Aber Musik ist nicht Sprache. Ihre Sprachähnlichkeit weist den Weg ins Innere, doch auch ins Vage. Wer Musik wörtlich als Sprache nimmt, den führt sie irre.

Sprachähnlich ist sie als zeitliche Folge artikulierter Laute, die mehr sind als bloß Laut. Sie sagen etwas, oft ein Menschliches. Sie sagen es desto nachdrücklicher, je höher die Musik geartet ist. Die Folge der Laute ist der Logik verwandt: es gibt Richtig und Falsch. Aber das Gesagte läßt von der Musik nicht sich ablösen. Sie bildet

kein System von Zeichen" (Adorno, 1956, zit. nach Knaus, 1973, S.71-72). "Der musikalische Inhalt aber ist in Wahrheit die Fülle alles dessen, was der musikalischen Grammatik und Syntax unterliegt" (Adorno, 1956, zit. nach Knaus, 1973, S.75).

Die Überzeugung von der Richtigkeit dieser These wird von zahlreichen Musikwissenschaftlern und Linguisten geteilt (Bright, 1963; Fay, 1971, 1974; Harweg, 1967, 1968, 1972; Kneif, 1973; Koch, 1971; Osmond-Smith, 1974; Nattiez, 1972, 1973; Nettl, 1958; Petri, 1966; Ruwet, 1972; Springer, 1956). Unsere Überzeugung von der Fruchtbarkeit der Analogie zwischen Sprache und Musik für die musikpsychologische Forschung bildet den Grundstein unserer Untersuchung: Wenn Sprache und Musik auf weite Strecken strukturell analog aufgebaut sind, dann könnte der strukturellen Analogie beider Medien eine funktionale Analogie der kognitiven Verarbeitung entsprechen. Damit sind auch die Konsequenzen für die musikalische Strukturbeschreibung offensichtlich: Wenn es gelingt, einen abgrenzbaren musikalischen Formtyp mit den methodischen Mitteln der modernen Linguistik zu analysieren und eine formale Syntax für diesen Formtyp zu formulieren, dann hätten wir eine Strukturbeschreibung an der Hand, die allen drei beschriebenen Funktionen gerecht werden könnte.

Musik besitzt - wie auch die Sprache - Strukturen, die nicht identisch sind mit dem, was man als "manifeste Reizstrukturen" bezeichnen kann, wie sie durch Tonhöhe, Klangfarbe, Lautstärke und ihrer Variation in der Zeit gegeben sind. Im Rahmen von Theorien und Experimenten zur auditiven Wahrnehmung wurden bislang nur diese manifesten Reizstrukturen mit Prozessen der kognitiven Verarbeitung in Beziehung gebracht. Aber schon einfache Melodien lassen über ihre physikalisch beschreibbare Struktur hinaus eine sich erst vom Melodieganzen und durch Vergleich einer Vielzahl von Melodien her erschließende S t r u k t u r

i n v a r i a n t e r B e z i e h u n g e n zwischen Einheiten verschiedener Größe und zeitlicher Reichweite erkennen. Diese Invarianz musikalischer Strukturen innerhalb eines Formtyps einer musikhistorischen Epoche stellt die Grundlage dafür dar, daß ein Hörer verschiedene musikalische Stile und Formen erkennen und voneinander differenzieren kann, wenn er dazu in der Lage ist, aufgrund seines Wissens solche Invarianzen zu erkennen.

Zur "psychologischen Realität" musikalisch-syntaktischer Strukturen

Unsere Untersuchung knüpft lose an ein psycholinguistisches Forschungsprogramm an und überträgt dieses auf den Bereich der Musikwahrnehmung, das unter dem Titel "Die psychologische Realität der Grammatik" bekannt geworden ist und für ein Jahrzehnt (von 1962 bis ca. 1972) die Hauptströmung der psycholinguistischen Forschung bildete.

Die Übertragung dieses Forschungsprogramms auf den Bereich der Musikwahrnehmung wurde wesentlich dadurch angeregt, daß es einigen Linguisten, die häufig zugleich Musikwissenschaftler sind, seit Anfang der siebziger Jahre immer besser gelang, musikalische Strukturen mit den formalen Mitteln der modernen Linguistik zu beschreiben (vgl. Stoffer & Boroda, 1980). So lag es nahe zu fragen, ob ein musikalisch versierter Hörer nicht mehr an kognitiver Aktivität beim Hören von Musik investiert, als die Theorien zur auditiven Wahrnehmung bislang zu erkennen geben. Das nämlich, was in Musik strukturell invariant ist und von einer musikalischen Syntax formal beschrieben wird, könnte für den Hörer ein Wissen sein, das er in Form spezifischer Erwartungen und anderer kognitiver Fertigkeiten beim Hören von Musik einsetzt. Hierdurch eröffnet sich für die Psychologie der auditiven Wahrnehmung eine Perspektive, die vielleicht mit der vergleichbar ist, die einige Sprachpsycho-

logen Ende der fünfziger Jahre sahen, nachdem sie Chomsky's (1957) Modell der generativen Grammatik rezipiert hatten.

Als diese erste Version der generativen Grammatik erschien, übte sie sogleich eine starke Faszination auf die zeitgenössischen Sprachpsychologen aus, denn was dort mit den formalen Methoden der Linguistik beschrieben wurde, nämlich die sprachliche Kompetenz eines idealen Sprecher/Hörers, las sich teilweise wie eine formalisierte psychologische Theorie der Sprachverwendung durch den menschlichen Geist. G.A. Miller sah sich hierdurch veranlaßt, in seinem programmatischen Aufsatz (Miller, 1962) zu bemerken, er glaube nun daran, daß der menschliche Geist auch für die Psychologie wieder eine Realität geworden sei. In Abkehr von den damals dominierenden Konditionierungs- und Assoziationstheorien entwarf er ein psycholinguistisches Forschungsprogramm, dessen Ziel im Nachweis der "psychologischen Realität" der von Chomsky vorgelegten Methode zur Generierung von Sätzen bestand.

Die später als Korrespondenz- oder Korrelationshypothese bezeichnete Vorstellung, daß die linguistischen Methoden der Satzgenerierung und Satzanalyse den psychologischen Prozessen der Produktion und Analyse von Sprache entsprechen könnten, führte in der Folgezeit zu dem Versuch, den einzelnen Schritten in der linguistischen Ableitung eines Satzes eine entsprechende Abfolge funktional analoger kognitiver Prozesse zuzuordnen, die eine Person beim Sprechen oder Hören ausführt. Insbesondere das linguistische Konzept der G e n e r i e r u n g von Sprache durch Anwendung von Regeln wurde wörtlich in die psychologische Theorie der sprachlichen Performanz übernommen.

So attraktiv der linguistische Ansatz sein mag, so naiv erscheint

in der Retrospektive die Gleichsetzung von linguistischer Kompetenz und sprachlicher Performanz. Spätere Arbeiten haben dann auch die Unhaltbarkeit der Korrelationshypothese mit zahlreichen Argumenten aufgezeigt (vgl. Hörmann, 1976).

Eine schwächere Version dieser Hypothese sieht in der Performanz lediglich einen "defizitären Modus des Kompetenzgebrauchs" (Engelkamp, 1974). "Psychologische Realität" beansprucht die Grammatik Chomsky's als Modell des beim Sprecher/Hörer vorhandenen Wissens über Sprache weiterhin, nicht dagegen die spezifischen Operationen, mit denen die Grammatik Sätze ableitet.

Legt man diese Hypothese zugrunde, dann offenbaren Untersuchungen zur "psychologischen Realität" der Grammatik die kognitive Organisation des grammatischen Wissens, nicht aber die Funktionsweise der mit diesem Wissen operierenden kognitiven Prozesse.

Die heute der psycholinguistischen Forschung zugrundeliegende schwächste Formulierung der Hypothese über die Beziehung zwischen linguistischem Modell und psychologischer Theorie der Sprachproduktion und des Verstehens von Sprache besagt, daß eine Grammatiktheorie nur ein System von Hypothesen über das sprachliche Wissen und seine Verwendung durch den Sprachbenutzer sein kann (Engelkamp, 1974). Die Frage nach einer "psychologischen Realität" einer Grammatik stellt sich so überhaupt nicht mehr, vielmehr tritt die Frage nach der psychologischen Brauchbarkeit von Hypothesen und Konstrukten über die Funktionen syntaktischen Wissens und die Form seiner kognitiven Repräsentation in den Mittelpunkt (Hörmann, 1976), die unabhängig von der linguistischen Brauchbarkeit einer Grammatik zu beurteilen ist.

Das Verhältnis von Linguistik und Psychologie wird so von der Einseitigkeit der Fragestellung entlastet, die in der Suche nach der "psychologischen Realität" zum Programm erhoben wurde, und eröffnet damit die für die Linguistik interessante Frage nach der linguistischen Brauchbarkeit psychologisch brauchbarer Hypothesen und Konstrukte einer performativen Theorie der Sprache.

Unser primäres Anliegen ist nun nicht die Formulierung einer nach linguistisch-musikwissenschaftlichen Kriterien adäquaten Syntax für einen bestimmten musikalischen Formtyp (obwohl wir uns auch dieser Aufgabe haben stellen müssen, vgl. Anhang 1), sondern die Formulierung eines nach psychologischen Gesichtspunkten adäquaten Modells der kognitiven Verarbeitung musikalischer Strukturen, in dem einer musikalischen Syntax die Rolle eines Hypothesengerüsts zukommt, das im Rahmen des strukturellen Teils eines solchen Modells Inhalt und Organisation der Repräsentation musikalischer Strukturaspekte und des syntaktischen Wissens des Hörers im Sinne einer "mentalen Grammatik" (Watt, 1973) beschreibt.

Plan der Untersuchung

Wir haben bis hierher ein Forschungsprogramm entwickelt, das unter zwei verschiedenen Perspektiven im Sinne von Herrmann (1976) sowohl als Teil eines Typ-a- als auch Teil eines Typ-b-Programms betrachtet werden kann:

"Wir unterscheiden a k z e n t u i e r e n d und in i d e a l i s i e r e n d e r Weise zwei Typen psychologischer Forschungsprogramme: T y p a, bei dem e i n e m (invarianten) empirischen, problematisierten Tatbestandsbereich (einer "domain" im Sinne von

Shapere (1974)) eine Serie von Theorien (als Sequenz versuchter Antworten auf eine invariante Domain-Problematisierung) gegenübersteht; Typ b, bei dem eine quasiparadigmatische Theoriekonzeption sozusagen auf eine Sequenz empirischer Tatbestände bzw. Probleme angewandt wird" (Herrmann, 1976, S.29).

Um ein Typ-a-Programm handelt es sich, wenn wir nach einer theoretischen Erklärung der kognitiven Verarbeitung beim Hören von Musik fragen. Hier knüpfen wir an bereits vorliegende Untersuchungen und theoretische Vorstellungen an, die auf dem Hintergrund der Psychologie der auditiven Informationsverarbeitung (vgl. Massaro, 1975, 1976) entstanden sind. Dabei handelt es sich um Untersuchungen zur Melodiewahrnehmung, allerdings unter den methodischen Einschränkungen einer Verwendung künstlicher Tonfolgen, die nicht den Anspruch erheben können, auch nur entfernte Verwandte musikalischer Melodien zu sein. Die theoretischen Unzulänglichkeiten des linearen Informationsverarbeitungsmodells bei der Erklärung der zutagegeförderten Phänomene verweisen auf die Notwendigkeit detaillierterer Annahmen über die Struktur von Tonfolgen bzw. Melodien und ihrer kognitiven Repräsentation. Von der hierarchischen Repräsentierbarkeit selbst einfachster und kürzester tonaler Strukturen ausgehend arbeiten wir die von Neumann (1980) und Wolff (1977) entwickelte theoretische Rahmenkonzeption eines strukturellen und funktionalen Hierarchiemodells der Informationsverarbeitung zu einem Modell des Strukturerkennens in Tonfolgen aus und überprüfen einige zentrale Hypothesen, die sich aus diesem Modell ableiten lassen, in zwei Experimenten.

Insofern als wir schon lange problematisierte Fragen der auditiven Wahrnehmung erneut aufgreifen und auf einem neuen theoretischen Hintergrund diskutieren, verfolgen wir ein Typ-a-Programm, das den Teil I unserer Untersuchung bildet.

Um ein Typ-b-Programm dagegen handelt es sich in Teil II unserer Untersuchung, wenn wir das skizzierte psycholinguistische Forschungsprogramm auf den Bereich der Musikwahrnehmung, speziell der Melodiewahrnehmung, übertragen. Hier fragen wir sowohl nach der Struktur der kognitiven Repräsentation musikalischer Strukturen, wobei das von uns entwickelte musikalische Syntaxmodell (Stoffer, 1979) als Hypothese über die Repräsentationsstruktur dient, als auch nach der Funktionsweise der Mechanismen kognitiver Verarbeitung beim Musikhören, wobei wir das in Teil I entwickelte Modell wieder aufgreifen und erweitern.

TEIL I

AUDITIVE MUSTERERKENNUNG: WAHRNEHMUNG UND REPRÄSENTATION TONALER STRUKTUREN

1. EINLEITUNG UND ÜBERBLICK

Gegenstand der theoretischen Diskussionen und experimentellen Arbeiten, die wir in Teil I unserer Untersuchung darstellen wollen, ist die Wahrnehmung von Tonfolgen, die nicht den Anspruch erheben, Musik zu sein. Dabei stehen diejenigen kognitiven Analyseprozesse im Mittelpunkt, die der Identifikation tonaler Strukturen dienen, also Identifikationsprozesse, die sich nicht ausschließlich auf einzelne Töne und deren individuelle Merkmale, sondern auf Beziehungsmuster zwischen Tönen (Intervallkonfigurationen, Konturverläufe) und ihre M e r k m a l e h ö h e r e r O r d n u n g (konfigurative Merkmale) beziehen.

Die Verwendung "unmusikalischer" Tonfolgen in einer Untersuchung, deren Anspruch es ist, strukturelles Hören von Musik zu erklären, bedarf einer besonderen Begründung, da es zunächst unplausibel erscheinen muß, Erkenntnisse über kognitive Prozesse des Musikhörens zu erwarten, wenn der Untersuchungsgegenstand mit Musik offenbar nicht viel zu tun hat.

Eine experimentelle Arbeit, deren Gegenstand die Wahrnehmung musikalischer Strukturen sein soll, erfordert eine systematische Variation der Parameter des "Reizmusters Musik" und als Voraussetzung dafür eine Definition aller innerhalb des Bereichs der Musik vorkommenden strukturellen Variationen, d.h. musikalische Strukturen müssen so beschrieben werden, daß überprüfbare Angaben über in Experimenten vorgenommene Strukturvariationen als unabhängige Variablen gemacht werden können. Hiermit verbunden ist also das Problem der O b j e k t i v i t ä t des experimentellen Vorgehens.

Um das Problem der experimentellen Objektivität zu lösen, ist es nicht ausreichend, ein allgemeines, formales Beschreibungssystem für b e l i e b i g e tonale Strukturen zu entwickeln, wie dies Jones (1974), Restle (197o) und Simon & Sumner (1968) mit bescheidenem Erfolg versucht haben, sondern es ist gleichzeitig eine Abgrenzung des "Objektbereichs Musik" von anderen Bereichen auditiv wahrnehmbarer Reize notwendig, um beurteilen zu können, ob eine vom Experimentator vorgenommene Strukturvariation noch eine musikalische ist oder bereits den Bereich der Musik verlassen hat, denn ohne ein Kriterium dafür, ob eine Tonfolge eine musikalische ist oder nicht, ist die über die Objektivität hinaus notwendige "externe Validität" (Campbell & Stanley, 1963) kaum zu gewährleisten.

Beim Versuch, eine Definition von Musik in Abgrenzung von anderen Schallereignissen zu geben, läßt sich selbst unter Musikwissenschaftlern nicht immer ein Konsensus erzielen, denn historische und zeitgenössische Erscheinungsformen von Musik haben oftmals nur so viel gemeinsam, daß es sich in allen Fällen um akustische Muster handelt, die auf den physikalisch definierten Dimensionen Frequenz und Schallintensität über die Zeit variieren. Moogs (1977) Versuch, die Musikpsychologie durch die Objekteigenschaften ihres Gegenstandes von anderen psychologischen Teildisziplinen abzugrenzen, mußte aus diesem Grunde von vornherein zum Scheitern verurteilt sein.

Da die Frage der Objektivität und externen Validität den psychologischen Experimentator offenbar vor unüberwindbare Schwierigkeiten stellt, könnte man versucht sein, eine Bearbeitung eines derart widerständigen Problemfeldes erst gar nicht in Betracht zu ziehen. Diese resignative Haltung kennzeichnet bis heute die Forschung im Bereich der Musikpsychologie.

Einen Ausweg bietet allerdings ein theoretisches Postulat: Wenn Musikhören psychologisch als Abfolge elementarer Verarbeitungsschritte zu erklären wäre, die unabhängig davon, ob es sich um musikalische oder willkürlich bzw. den experimentellen Notwendigkeiten entsprechend konstruierte Tonfolgen handelt, ihre Funktionscharakteristika und ihre Abfolge nicht ändern und wenn man weiterhin annehmen könnte, daß beim Hören von Musik im Vergleich zum Hören künstlich konstruierter Tonfolgen keine weiteren Verarbeitungsschritte hinzukommen, dann würde die Frage nach den spezifisch musikalischen Strukturen in diesem Zusammenhang gegenstandslos. Akzeptiert man dieses Postulat, dann kann man als Experimentator ohne Skrupel die physikalischen Dimensionen Frequenz und Schallintensität unabhängig voneinander systematisch in der Zeit variieren, ohne das Risiko eingehen zu müssen, den psychologischen Forschungsgegenstand "Musikhören" aus den Augen zu verlieren.

Was wir bis heute über "Musikwahrnehmung" wissen, ist offensichtlich auf dem Hintergrund dieses oder eines ähnlich restriktiven Postulats entstanden. Der diesen Untersuchungen zugrundeliegende kognitionspsychologische Ansatz ist der der "Psychologie der Informationsverarbeitung", speziell des "linearen Informationsverarbeitungsmodells" (z.B. Massaro, 1975). Auf der Basis dieses Ansatzes wurden in den vergangenen zehn Jahren immense Fortschritte im Verständnis menschlicher Informationsverarbeitungsprozesse erzielt, die mit gewissen Einschränkungen auch dazu geeignet sind, Fragen zur kognitiven Verarbeitung von Musik zu beantworten.

Die Einschränkungen resultieren aus der Tatsache, daß in den allerwenigsten Fällen echte Musik verwendet wurde, und wenn dies doch einmal

der Fall war, dann mehr zum Zwecke der Demonstration eines hübschen Effekts als im Rahmen einer funktionalen Analyse kognitiver Prozesse des Musikhörens. Damit wollen wir keineswegs behaupten, daß alle vorliegenden Befunde zur auditiven Wahrnehmung künstlicher Tonfolgen im Zusammenhang der Erklärung bestimmter Aspekte der Musikwahrnehmung wertlos seien. Dies kann schon deshalb nicht der Fall sein, weil es unwahrscheinlich ist, daß das auditive System spezielle Verarbeitungskomponenten entwickelt hat, die ausschließlich der Analyse musikalischer Strukturen dienen. Vielmehr muß man davon ausgehen, daß die kognitive Verarbeitung künstlicher Tonfolgen bis auf weite Strecken mit der von Musik übereinstimmt; man muß sich allerdings auch der Frage stellen, wie weit diese angenommene Übereinstimmung eigentlich geht und welche kognitiven Verarbeitungsmechanismen beim Hören von Musik d a r ü b e r h i n a u s wirksam werden können. Spätestens aber dann ist es notwendig, dem Spezifikum musikalischer Strukturen mehr als bisher Aufmerksamkeit zu schenken.

Den Ausgangspunkt unserer Untersuchung bilden die bisher vorhandenen Theorien und experimentellen Befunde zum Erkennen von Strukturen in künstlich konstruierten Tonfolgen, die bis auf wenige Ausnahmen auf dem Hintergrund des linearen Informationsverarbeitungsmodells gewonnen wurden. Wir werden dabei feststellen, daß das lineare Modell bereits bei der Erklärung der Strukturidentifikation in künstlichen Tonfolgen auf große Schwierigkeiten stößt, die auf die Notwendigkeit verweisen, tonale Strukturen (Strukturen künstlicher wie musikalischer Tonfolgen) so zu beschreiben, daß diese Beschreibung eine Hypothese über die Struktur der kognitiven Repräsentationen darstellt. Experiment 1 demonstriert unter Verwendung künstlich konstruierter Tonfolgen die Unhaltbarkeit der Linearitätsannahme des Informationsverarbeitungsmodells und liefert gleichzeitig Befunde, die mit einer alternativen

theoretischen Konzeption vereinbar sind.

Die engere Fragestellung von Experiment 1 zielt auf im Rahmen des linearen Ansatzes ungelöst gebliebene Probleme der auditiven Wahrnehmung von Strukturen in Tonfolgen, nämlich auf das der Erklärung unterschiedlicher Schwierigkeiten der Identifikation eines Tonhöhenverlaufsmusters in Abhängigkeit von der Struktur der Tonfolgen und auf das der Erklärung der Ganzheitlichkeit der Repräsentation von Tonfolgen. Unser Modell des Erkennens von Strukturen in Tonfolgen erweist sich als teilweise adäquate Erklärung der Befundmuster. Ein überraschendes Ergebnis zwingt uns dann jedoch zu einer weiteren theoretischen Präzisierung, deren Adäquatheit sich bei der Interpretation der Befunde von Experiment 2 herausstellt.

Das in Teil I entwickelte Modell beschränkt seine Aussagen keineswegs nur auf derart kurze und künstliche Tonfolgen, wie wir sie in den beiden ersten Experimenten verwendet haben. Die theoretische Konzeption wird von Anfang an unter dem Gesichtspunkt entwickelt, die kognitive Verarbeitung musikalischer Strukturen zu modellieren. Inwieweit dies gelungen ist und welche weiteren theoretischen Präzisierungen notwendig sind, um nicht nur die Wahrnehmung künstlich konstruierter Tonfolgen erklären zu können, ist dann Gegenstand der Diskussion und Experimente von Teil II.

2. AUDITIVE INFORMATIONSVERARBEITUNG DER STRUKTUR VON TONFOLGEN

Wir wollen zunächst Theorien der auditiven Informationsverarbeitung speziell unter dem Aspekt der Erklärung auditiver Mustererkennungsleistungen diskutieren. Der Terminus "Mustererkennung" stammt aus dem Bereich der Computertechnologie (vgl. Neisser, 1967) und bezieht sich dort auf das maschinelle Erkennen von Zeichen durch Automaten. Die bei der Programmierung und Systementwicklung von Computern gewonnenen Erkenntnisse über die notwendigen Verarbeitungsschritte, die innerhalb eines automatischen Zeichenerkennungssystems ablaufen müssen, wurden einerseits auf das menschliche Erkennen visueller Formen übertragen (vgl. Reed, 1978; Shaw, 1973; Uhr, 1963; Uttal, 1975), und psychologische Erkenntnisse über das menschliche Formenerkennen befruchteten andererseits technische Lösungen dieses Problems.

Ursprünglich wurde der Begriff "Mustererkennung" ausschließlich mit den präattentiven Prozessen der Merkmalsanalyse in Zusammenhang gebracht (vgl. Neisser, 1967), mittlerweile wird er aber im Bereich der "artificial intelligence" Forschung auch zur Bezeichnung komplexerer Identifikationsprozesse bei der visuellen Analyse realer Umweltszenen verwendet, für die hochstrukturierte Repräsentationssysteme und Selektionsprozesse angenommen werden müssen, die sich einer vereinfachenden Kategorisierung in präattentive und attentive Prozesse entziehen (vgl. Neumann, 198o), oder er wird durch den Begriff der "Szenenanalyse " ersetzt, um damit die größere Komplexität der zu erkennenden Reizmuster und der beteiligten kognitiven Prozesse gegenüber denen der Mustererkennung (in der ursprünglichen Bedeutung dieses

Begriffs) hervorzuheben.

Unter dem Begriff "auditive Mustererkennung" [1] fassen wir eine Reihe kognitiver Prozesse zusammen, die gemeinsam am Erkennen struktureller oder konfigurativer Merkmale akustischer Ereignisfolgen beteiligt sind. Dabei verstehen wir unter strukturellen oder konfigurativen Merkmalen solche, die durch Merkmalsvariationen in der Zeit entstehen ("emergent properties" oder "configurational properties", vgl. Garner, 1978), im Gegensatz zu statischen Merkmalen, die an einem bestimmten Ereignis zu einem bestimmten Zeitpunkt wahrzunehmen sind. Zur Mustererkennung gehören so relativ "einfache" Wahrnehmungsprozesse, wie das Erkennen der Klangfarben von Instrumenten, wobei auch zeitliche Charakteristika eine bedeutende Rolle spielen (Risset & Matthews, 1969), sowie so komplexe Prozesse, wie das Erkennen stilistischer Wandlungen in den Kompositionen des jungen im Vergleich zum alten Beethoven. Im Zusammenhang mit unserem Thema stehen solche Mustererkennungsprozesse im Mittelpunkt, die sich auf das Erkennen von Tonhöhen- und Intervallrelationen in der Zeit beziehen, d.h. wir beschränken uns vorläufig auf den horizontalen Strukturaspekt von Musik und vernachlässigen die vertikalen Schichten (z.B. Harmonien, Kontrapunkt und andere Formen von Beziehungen zwischen verschiedenen Stimmen im homophonen und polyphonen Satz).

Im Rahmen der folgenden Diskussion von Theorien und Befunden zur auditiven Mustererkennung in Tonfolgen stellen wir die Frage, wie diese

---

1) Wir hätten auch den neuen Begriff "auditive Strukturerkennung" prägen können, verwenden aber lieber den der "auditiven Mustererkennung", um hiermit auf die funktionalen und strukturellen Analogien zur visuellen Mustererkennung zu verweisen.

Theorien das Erkennen von Tonhöhen- und Intervallrelationen und deren Variation in der Zeit erklären können und ob diese Erklärungsansätze auch geeignet sind, sehr komplexe Formen der musikalischen Mustererkennung zu erklären, deren Komplexitätsgrad an unser Beispiel des Erkennens stilistischer Wandlungen heranreicht. Die gemeinsame Klammer für diese Theorien stellt der Informationsverarbeitungsansatz dar. Im Verlauf der Diskussion werden sich einige (bisher wenig beachtete) Problembereiche herauskristallisieren, die für eine kognitive Theorie der Musikwahrnehmung von zentraler Bedeutung sind. Diese Problembereiche greifen wir dann in den Fragestellungen der Experimente von Teil I auf.

2.1 AUDITIVE INFORMATIONSVERARBEITUNG: CHARAKTERISIERUNG DES ANSATZES

Mit Erscheinen von Neisser's Monographie "Cognitive Psychology" (1967) wurde ein "Paradigmenwechsel" (Kuhn, 1962) endgültig vollzogen, der sich bereits ein Jahrzehnt vorher angekündigt hatte, als die ersten explizit formulierten Informationsverarbeitungsmodelle publiziert wurden. Zu den Pionieren auf diesem Gebiet gehören insbesondere Broadbent (1958), Haber (1967), Miller, Galanter & Pribram (1960), Sperling (1963) sowie Sternberg (1966, 1967).

An der Wiege der Informationsverarbeitungsmodelle standen häufig technische Konzepte Pate, die parallel zur fortschreitenden technischen Entwicklung zu immer komplexeren Informationsflußdiagrammen als Modelle menschlichen Erkennens führten: Die Entwicklung begann mit Analogien zu technischen Nachrichtenübermittlungssystemen mit "Filtern" und "Kanälen" (Broadbent, 1958), wurde dann durch Analogien zu selbst-

regulierenden kybernetischen Systemen mit "Regelkreisen" ergänzt (Miller, Galanter & Pribram, 196o) und mündete schließlich in Analogien zur sophistiziertesten Maschine, dem Computer und seiner Software (Neisser, 1967; Pribram, Nuwer & Baron, 1974).

Dieses jüngste Kind der Psychologie wurde auf den Namen "Kognitive Psychologie" getauft. Das in der Namensgebung durch Neisser (1967) enthaltene Programm des Informationsverarbeitungsansatzes kommt in seinem vielzitierten Manifest zum Ausdruck:

"As used here, the term 'cognition' refers to all processes by which the sensory input is transformed, reduced, elaborated, stored, recovered, and used. It is concerned with these processes even when they operate in the absence of relevant stimulation, as in images and hallucinations. Such terms as sensation, perception, imagery, retention, recall, problem solving, and thinking, among others, refer to hypothetical stages or aspects of cognition" (Neisser, 1967, S. 4).

Die Attraktivität der Computeranalogie rührt daher, daß ein Computer einiges von dem zu leisten vermag, was man vor dem Computerzeitalter als spezifisch menschliche Leistungen betrachtete. Mit diesen Analogien glaubte man, den Realitätsanspruch geistiger Prozesse wieder restaurieren zu können, der durch den Behaviorismus abhandengekommen war.

Analog der Unterscheidung zwischen Hardware und Software eines Computers wird in Informationsverarbeitungsmodellen zwischen der Abfolge von Verarbeitungsstufen und den ihnen zugeordneten Speichern (Struktureller Aspekt) und den Verarbeitungsprozessen (funktionaler Aspekt) differenziert (Garner, 198o). Zwei Extreme bezüglich der Anzahl postulierter Speicher- und Verarbeitungsstufen stellen das vierstufige Modell von Massaro (1975, 1976) und das einstufige Modell von Craik

& Lockhart (1972; Lockhart, Craik & Jacoby, 1976) dar. Zur Kritik des Craik & Lockhart-Modells vgl. Baddeley (1978) und Eysenck (1978).

Eine grundlegende Voraussetzung dieser Modelle ist die L i n e a r i t ä t der Abfolge der Informationsverarbeitungsstufen von peripheren, sensorischen Prozessen bis hin zur Herstellung von abstrakten, semantischen Repräsentationen. Mit der Linearitätsannahme gekoppelt ist die der aktiven K o n s t r u k t i o n des Wissens über die Welt aus e l e m e n t a r e n B r u c h s t ü c k e n unter Verwendung vonGedächtnisbeständen.

Gregory (1972) ist einer der wenigen, die die konstruktivistische Annahme auch für die Wahrnehmung explizit formulieren:

"... perceptions are constructed, by complex brain processes, from fleeting fragmentary scraps of data signalled by the senses and drawn from the brain's memory banks - themselves constructions from snippets of the past, normal everyday perceptions are not selections of reality but rather imaginative constructions - fictions - based ... more on stored past than on the present" (Gregory, 1972, S. 7o7, zit. nach Mace, 1974, S. 14o).

Das erste Informationsverarbeitungsmodell, das explizit die Wahrnehmung von Musik thematisiert, wurde von Deutsch (1969) beschrieben. Deutsch (1978) charakterisiert den Ansatz der musikalischen Informationsverarbeitung ähnlich, wie Gregory ihn für die visuelle Modalität beschrieben hat:

"The a b s t r a c t i o n of tonal information occurs in several s t a g e s . The initial stages of abstraction involve detection of both specific features (such as intervals, chords, and tone chroma) and global features (such as contour). Such features are then c o m b i n e to form tonal sequences, which are themselves abstracted and combined in systematic ways" (Deutsch, 1978, S. 192, meine Hervorhebungen).

Creel, Boomsliter & Powers (197o) ergänzen diese Charakterisierung durch die Einbeziehung der Funktion von Gedächtnisprozessen bei der Wahrnehmung von Tonfolgen:

"What we experience is a form we i m p o s e . Its raw material is m e m o r y of the past portion of the stimulus and expectation of its future. The auditory input itself is, at any given instant, a s i n g l e s t i m u l u s , which, studing alone, has no pattern or meaning" (Creel et al., 197o, S. 534, meine Hervorhebungen).

Den auditiven Informationsverarbeitungsmodellen liegen also dieselben Annahmen zugrunde, wie denjenigen für die visuelle Modalität: Die Verarbeitung des akustischen Reizmusters "Musik" wird als eine Sukzession von aufeinander aufbauenden Prozessen betrachtet, die über mehrere Verarbeitungsstufen unter Einbeziehung von Gedächtnisinformationen aus elementaren Verarbeitungseinheiten die Wahrnehmung der Musik oder Melodie konstruieren.

Diese Voraussetzungen sollten im Prinzip geeignet sein, die Basis für eine Erklärung auditiver Mustererkennungsprozesse selbst bei komplexesten musikalischen Strukturen zu liefern, solange sich keine experimentellen Befunde einstellen, die mit der Linearitätsannahme für die Prozeßreihenfolge und mit der konstruktivistischen Hypothese über die Herstellung bewußter Repräsentationen des Wahrgenommenen aus isolierten, kleinsten Verarbeitungseinheiten nicht vereinbar sind.

Bei der folgenden Diskussion der Informationsverarbeitungsmodelle der auditiven Mustererkennung wird dieser Gesichtspunkt im Vordergrund stehen; gleichzeitig werden wir aber auch unser Augenmerk auf

widersprüchliche Erklärungsansätze verschiedener Autoren richten und eventuell vernachlässigte strukturelle und/oder funktionale Aspekte der kognitiven Verarbeitung von Musik herausarbeiten. Dabei orientieren wir uns nicht so sehr an dem eigentlich naheliegenden Ordnungsprinzip der Prozeßreihenfolge, sondern diskutieren Hypothesen über Prozeßreihenfolgen für die verschiedenen Typen von Mustererkennungsleistungen, die in der experimentellen Literatur beschrieben sind. Dieses Vorgehen hat den Vorteil, Aufgabenabhängigkeiten der Mustererkennungsprozesse und Zusammenhänge zwischen unterschiedlichen Verarbeitungsstufen deutlich werden zu lassen.

2.2 MUSTERERKENNUNGSLEISTUNGEN BEIM HÖREN VON TONFOLGEN UND DIE IHNEN ZUGRUNDELIEGENDEN PROZESSE DER INFORMATIONSVERARBEITUNG

Einen partiellen Überblick über Mustererkennungsleistungen beim Hören von Tonfolgen enthalten die Arbeiten von Davies (1978) und Deutsch (1978). (Sammelreferate über die hier nur am Rande berücksichtigte Wahrnehmung isolierter akustischer Signale geben Nordmark (1978) und Ward (197o) für die Tonhöhe, Scharf (1978) für die Lautstärke und Terhardt (1978) für komplexere Qualitäten einzelner und simultan erklingender Töne).

Aus den in der Literatur beschriebenen Mustererkennungsleistungen lassen sich drei Typen bilden: 1. Erkennen der Beziehungen zwischen Tönen, 2. Erkennen der Melodiekontur (Muster des Auf- und Absteigens der Intervalle) und 3. Organisation bei Wahrnehmung und Speicherung von Tonfolgen.

2.2.1 ERKENNEN DER BEZIEHUNGEN ZWISCHEN TÖNEN

Beim Musikhören dürfte für den Hörer nur äußerst selten eine Situation gegeben sein, in der er bewußt auf die Reihenfolge einzelner Töne achtet; meist konzentriert sich unsere Aufmerksamkeit auf größere Abschnitte von Tonhöhenverläufen, also auf das, was Metzger (1966) und Wellek (1963) "Zeitgestalt" nennen.

"In the end, it is the perception of overall form and structure which is important" (Davies, 1978, S. 47), aber an anderer Stelle schreibt Davies:

"One way of conceptualizing a piece of music is thus as a series of events of the present, each one succeeding the previous one in an orderly series..." (Davies, 1978, S. 62).

Wenn einerseits beim Hören von Musik nicht die Wahrnehmung der sie konstituierenden kleinsten Elemente, der Töne, phänomenal im Vordergrund zu stehen scheint, andererseits die Erklärung der Wahrnehmung von Musik im Rahmen des Informationsverarbeitungsansatzes beim einzelnen Ton ihren Ausgangspunkt nimmt, dann muß erklärt werden, durch welche kognitiven Integrationsleistungen der Eindruck von Beziehungen zwischen Tönen und größeren Tongruppen entsteht.

Ein methodischer Ansatz in Richtung auf dieses Erklärungsziel besteht in der Darbietung kurzer Tonfolgen, in denen die Parameter Tonhöhe und SOA [1] systematisch variiert werden, die ein Hörer ent-

---

[1] Stimulus onset asynchronity: Zeitintervall zwischen dem Beginn zweier sukzessiver Reize

weder hinsichtlich der Reihenfolge (meistens bei nur zwei Tönen, z.B. durch die Aufgabe zu beurteilen, ob zuerst ein hoher oder ein tiefer Ton auftrat) beurteilen soll oder - insbesondere bei längeren Tonfolgen - deren Verlaufsmuster er auf Kategorien von Verlaufsmustern abbilden soll, indem er das wahrgenommene Muster von n bekannten Alternativen diskriminiert. Die Variationen des Schwierigkeitsgrades dieser Aufgaben (gemessen durch die relativen Fehlerhäufigkeiten) in Abhängigkeit von den variierten Parametern der Tonfolgen erlauben dann Rückschlüsse auf die Funktionsweise der bei der Aufgabenbewältigung beteiligten kognitiven Prozesse.

Die mit diesem methodischen Ansatz häufig verbundene Erklärungsstrategie besteht darin, aus extrem hohen Fehlerraten auf Begrenzungen des kognitiven Systems zu schließen und in der Umkehrung dieses Schlusses eine Abnahme der Fehlerrate als eine systemadäquate kognitive Verarbeitung zu interpretieren. Durch die Kenntnis dessen, was ein System nicht zu leisten vermag, kann die Zahl möglicher Modellalternativen zur Erklärung der dem System zugrundeliegenden Prozesse und Strukturen stark eingeschränkt werden.

Erkennen der zeitlichen Reihenfolge in Abhängigkeit vom SOA

Eine Form der Begrenzung der auditiven Mustererkennungsprozesse besteht in der Größe des minimalen SOAs, das zwischen dem Beginn zweier sukzessiver Töne verstreichen muß, um ein Urteil über die Reihenfolge der Töne zu ermöglichen. (Einer solchen Aufgabe sieht sich z.B. ein Dirigent gegenüber, der der Blechbläsergruppe

seines Orchesters den Einsatz gibt und dabei bemerkt, daß die Tuba zu früh eingesetzt hat.)

Die Untersuchung der Frage, welches das minimale SOA ist, das ein Erkennen der Reihenfolge zweier Töne gestattet, begann mit den Arbeiten von Hirsh (Hirsh, 1959, 1974; Hirsh & Fraisse, 1964; Hirsh & Sherrick, 1961). Hirsh (1959) bot seinen Vpn zwei Töne mit variablem SOA dar, die nach 5oo msec gemeinsam ausklangen. Die Vpn hatten jeweils zu beurteilen, ob der höhere oder der tiefere Ton zuerst erklungen war. Um eine durchschnittliche Erkennungsrate von 75 % richtiger Urteile zu erreichen, mußte das SOA ca. 2o msec betragen. Eine Wiederholung dieser Experimente auch in anderen Sinnesmodalitäten (Hirsh & Sherrick, 1961) zeigte, daß das SOA, das ein E r k e n n e n d e r z e i t l i c h e n S u k z e s s i o n ermöglicht, mit ca. 1 bis 2 msec (vgl. auch Green, 1971) für die auditive Modalität kleiner als für alle anderen Modalitäten ist und zwischen den einzelnen Modalitäten stark schwankt, während das E r k e n n e n d e r R e i h e n f o l g e zweier Reize für alle Modalitäten konstant bei ca. 2o msec SOA liegt, was auf einen zentralen Mechanismus für das Erkennen zeitlicher Reihenfolgen schließen läßt (Peters, 1975).

Bei einem SOA von nur 2o msec bilden die beiden Töne noch eine gemeinsame perzeptive Einheit, d.h. sie sind noch nicht als zeitlich getrennte und eigenständige Töne mit differenzierenden Merkmalen wahrnehmbar. Dies liegt einmal an der bis ca. 15 msec nach Einsetzen eines Tones dauernden Energieintegration, die durch Summierung der Amplitude über diesen Zeitraum den Signal-Geräuschspannungsabstand erhöht (Doughty & Garner, 1947; Green, 1973; Stevens & Davies, 1938; Zwislocki, 196o), was phänomenal durch einen Wechsel

von einem Klickgeräusch unbestimmter Tonhöhe bei Tondauern unterhalb von 15 msec zu einem Ton definierter Höhe bei Tondauern oberhalb von 15 msec angezeigt wird, und zum anderen an der für SOAs zwischen 15 und ca. 6o bis 8o msec nachweisbaren Merkmalsintegration, wie sie sich in Experimenten zur Vorwärtsmaskierung und zur rückwirkenden Maskierung zeigt (einen umfassenden Überblick über die Literatur zur auditiven Maskierung geben Massaro, 1972, 1976a; Patterson & Green, 1978 und Zwislocki, 1978).

Der Befund von Hirsh (1959) ist also nicht so zu interpretieren, daß das Reihenfolgeurteil durch ein Erkennen der einzelnen Tonhöhen und deren Vergleich ermöglicht wurde. Nickerson & Freeman (1974) führen das Erkennen der Reihenfolge auf einen Prozeß zurück, den sie als "Spektralanalyse" bezeichnen: Bei Einsetzen des zweiten Tons wird der weiterhin erklingende erste Ton mit dem zweiten integriert (durch Merkmalsintegration), wodurch ein Klangfarbenwechsel entsteht; ist nun der zweite Ton höher als der erste, hört man ein "blipp"-ähnliches Geräusch, ist der zweite Ton tiefer als der erste, hört man dagegen ein "blapp"-ähnliches Geräusch. Aufgrund der im Versuch gegebenen Rückmeldungen können die Vpn die Diskrimination und die richtige Zuordnung von Klangfarbe und Reihenfolge lernen. Durch ein Training in der Diskrimination und Kategorisierung von Klangfarbenwechseln läßt sich nicht nur die Reihenfolge von zwei Tönen bei SOAs kleiner als 8o msec erschließen (Efron, 1973; Patterson & Green, 197o), sondern auch die von bedeutend längeren Tonfolgen (Warren, 1974a).

Die Wahrnehmung einer kontinuierlichen Tonhöhenänderung bei einem Glissando ist auf die Integration der Tonhöhen durch Merkmalsintegration zurückzuführen. Wenn die Reihenfolge von Tönen zu beurteilen

ist, wird diese Aufgabe durch Einfügen von Glissandi in der Regel erschwert (Kerivan & Carey, 1976), andererseits erleichtern Glissandoübergänge zwischen Tönen, die ein großes Intervall bilden, das Erkennen der Reihenfolge der Töne im Vergleich zur selben Tonfolge mit Pausen statt Glissandi (Bregman & Dannenbring, 1973).

Im Widerspruch zu dem von Hirsh (1959) und Hirsh & Sherrick (1961) angegebenen minimalen SOA von 2o msec für das Erkennen der Reihenfolge von Tönen scheinen die Befunde von Broadbent & Ladefoged (1959) zu stehen. Im Unterschied zu Hirsh verwendeten die Autoren qualitativ unterschiedliche Geräusche (Zischen, Piepen und Summen) und boten je zwei dieser Geräusche mit einem SOA von 15o msec dar, wobei sich eine Fehlerrate von ca. 5o % ergab. Broadbent & Ladefoged (1959) führen den Unterschied zu den Befunden von Hirsh auf das Training der Vpn Hirsh's zurück, aber Grengel & Hirsh (197o, zit. bei Hirsh (1974) ohne Quelle) haben gezeigt, daß untrainierte Vpn bei denselben Reizen, wie Hirsh sie verwandte, von einem SOA größer als 45 msec an in 75 % der Fälle die Reihenfolge richtig angeben können.

Warren, Obusek, Farmer & Warren (1969) sowie Thomas & Fitzgibbons (1971) ermittelten sogar noch viel größere minimale SOA-Werte mit qualitativ unterschiedlichen Geräuschen, die zyklisch wiederholt wurden: trainierte Vpn benötigen SOAs von wenigstens 3oo msec, untrainierte sogar solche von ca. 7oo msec, um die richtige zeitliche Reihenfolge zu erkennen; bei einmaliger Darbietung fiel das kritische SOA auf ca. 2oo msec ab (Warren, 1972). Deutlich niedrigere Werte bei hohem Trainingsniveau berichten Neisser & Hirst (1974) und Nickerson & Freeman (1974).

Bei Vergleich der geschilderten empirischen Sachverhalte ergeben sich vier methodische Gesichtspunkte, auf die die Befunddiskrepanzen eventuell zurückgeführt werden können: 1. Qualitative Unterschiede zwischen den verwendeten Aufgaben (Diskrimination vs. absolute Identifikation); 2. zyklische Wiederholung des Reizmusters vs. einmalige Darbietung; 3. Anzahl der Signale eines Reizmusters und 4. Unterschiede im Trainingsniveau der Vpn.

Zu 1.: Eine Diskrimination zwischen zwei qualitativ unterschiedlichen Klangfarben, wie sie bei Hirsh (1959) verlangt wurde, braucht nicht notwendigerweise auf ein Reihenfolgeurteil zurückzuführen sein (Peters, 1975). Die Reihenfolge zweier Signale zu beurteilen heißt doch, die Merkmale der Signale getrennt und in zeitlicher Sukzession zu erkennen, was aber bei integrierten Tönen, die phänomenal nicht den Status individueller Ereignisse besitzen, nicht der Fall sein kann. Die Diskriminationsleistung bei Hirsh (1959) bezieht sich auf ein Erkennen eines qualitativen Wechsels und die Eindeutigkeit seiner Zuordnung zu den beiden Reaktionsalternativen, während sich das Erkennen der Reihenfolge in den Untersuchungen von Warren et al. (1969), Warren (1972), Neisser & Hirst (1974) sowie Nickerson & Freeman (1974) auf eine absolute Identifikation der das Reizmuster konstituierenden Ereignisse bezieht.

Zu 2.: Bei zyklisch wiederholter Darbietung der Reizfolge wurden stets höhere kritische SOA-Werte ermittelt als bei einmaliger Darbietung (Peters & Wood, 1973a, b; Warren, 1972). Die Erhöhung der Schwierigkeit der Mustererkennung ist bei zyklischer Darbietung wahrscheinlich auf die Uneindeutigkeit von Anfang und Ende und der sich daraus ergebenden unterschiedlichen s u b j e k t i v e n

O r g a n i s a t i o n des Reizmusters bei mehreren Wiederholungen zurückzuführen (vgl. Royer & Garner, 197o).

Zu 3.: Wird das Reizmuster nur einmal dargeboten, zeigt sich ein deutlicher Effekt der Elementzahl: bei vier Elementen fand Warren (1972) ein kritisches SOA von 2oo msec, bei drei Elementen ermittelten Peters & Wood (1973a, b) ein SOA von nur noch 5o msec und für nur zwei Elemente ergaben Hirsh's (1959) Untersuchungen ein kritisches SOA von 2o msec. Je höher aber die Anzahl der ein Muster bildenden Elemente, desto größer ist die Zahl der alternativen Muster, die durch Permutation der Elemente entstehen können und zwischen denen die Vpn bei der Vorbereitung ihrer Reaktion unterscheiden müssen. Der Zusammenhang zwischen Alternativenzahl und Aufgabenschwierigkeit wird durch das Hick'sche Gesetz (Hick, 1952) quantitativ beschrieben, wobei dort allerdings die Aufgabenschwierigkeit durch die Reaktionszeit (RT) operationalisiert ist. Diese Beziehung wird im Rahmen des Informationsverarbeitungsansatzes funktional durch eine Zunahme der Beanspruchung von zentraler Verarbeitungskapazität erklärt (vgl. Sanders, 1971).

Eine ähnliche und hiermit vielleicht zusammenhängende Beobachtung wurde von Garner & Clement (1963) sowie Royer & Garner (1966, 197o) gemacht: Die Schwierigkeit des Erkennens eines visuellen Musters oder einer Tonfolge hängt von der Anzahl möglicher Organisationsalternativen ab. Da aber Tonfolgen mit wenigen Elementen auch weniger Möglichkeiten alternativer Organisationen besitzen, sollte das Erkennen ihrer Reihenfolge einfacher sein, wenn man annimmt, daß zum Erkennen hierbei weniger Vergleichsprozesse mit Gedächtnisrepräsentationen notwendig sind.

Wollte man diesen Zusammenhang verallgemeinern, käme man zu der mit der Phänomenologie im Widerspruch stehenden Hypothese, daß das Erkennen einer Melodie um so schwieriger sein sollte, je länger diese ist, und da die zentrale Verarbeitungskapazität begrenzt ist (z.B. Broadbent, 1958; Deutsch & Deutsch, 1963; Kahneman, 1973; Moray, 1969; Treisman, 1964), sollte von einer bestimmten Länge an überhaupt keine Mustererkennung für eine ganze Melodie mehr möglich sein. Wir werden weiter unten noch zeigen, daß durch bestimmte Hypothesen über die Organisation von Mustern und die daran beteiligten Gedächtnisprozesse auch eine Erklärung der Mustererkennung bei längeren Melodien möglich ist.

Zu 4.: Intensives Training kann die Schwierigkeiten bei der auditiven Mustererkennung so dramatisch reduzieren (Nickerson & Freeman, 1974; Neisser & Hirst, 1974), daß es nicht mehr sinnvoll erscheint, von einer Begrenzung der auditiven Mustererkennung durch die Geschwindigkeit der Signalfolge zu sprechen. Was im Verlauf eines solchen Trainings stattfindet, läßt sich wohl am besten mit Gibson & Gibson's (1955) Konzept des Wahrnehmungslernens beschreiben: Die bei sehr kurzen SOAs zu einer sich über die Zeit verändernden Klangfarbe integrierten Signale besitzen in Abhängigkeit von ihrer Reihenfolge unterschiedliche Verteilungen der Klangfarbenvariation über die Zeit, so daß eine zunehmend differenziertere Wahrnehmung dieser Verteilungen eine Diskrimination zwischen den verschiedenen Reihenfolgen erlaubt.

Eine Differenzierung der Wahrnehmung durch Lernen erklärt zwar die Fähigkeit des Hörers, integrierte Signalfolgen hinsichtlich der Reihenfolge ihrer Elemente zu diskriminieren, nicht dagegen

die Schwierigkeiten der Mustererkennung oberhalb des SOA-Bereichs, von dem ab einzelne, voneinander abgehobene Töne erkennbar sind.

Erklärungen für die differentiellen Schwierigkeiten der auditiven Mustererkennung bei verschiedenen SOAs haben Neisser & Hirst (1974) und Warren (1974a, b; Warren & Ackroff, 1976) durch die Annahme zweier funktional unterschiedlicher Prozesse der Mustererkennung gegeben.

Neisser & Hirst's (1974) Erklärung besteht darin, zwei Formen der Repräsentation eines akustischen Musters zu postulieren: die erste Form, die analoge Repräsentation ("analog"), besteht in diskriminativen Merkmalen des Reizmusters; die zweite Form der Repräsentation wird auf der Grundlage der ersten konstruiert und besteht in der Repräsentation einer zeitlich geordneten Folge der das Muster bildenden Elemente ("strings"). Die Konstruktion eines Strings wird als eine Sequenz von Einzelentscheidungen über die Reihenfolge von je zwei Elementen aufgefaßt, so daß die Herstellung einer Stringrepräsentation um so länger dauern muß, je mehr Elemente das Muster enthält. Bei großen SOAs kann nun der Hörer die Stringrepräsentation mit derselben Geschwindigkeit herstellen, mit der die Tonfolge dargeboten wird, d.h. Wahrnehmung der Tonfolge und Kodierung des Strings verlaufen zeitlich parallel. Wenn aber die SOAs kleiner als die zur Kodierung eines Elements der Stringrepräsentation benötigte Zeit sind, dann kann die Konstruktion der Stringrepräsentation erst n a c h Ende der Tonfolge beginnen. Die Übungseffekte lassen sich z.T. dadurch erklären, daß die Zeit, die bei der Konstruktion des Strings für ein Element benötigt wird, mit zunehmendem Übungsgrad kürzer wird, so daß der SOA-Bereich nach unten ausgedehnt wird, indem eine Parallelität zwischen Wahrnehmung und Konstruktion der Stringrepräsentation

vorliegt. Aber bei SOAs unterhalb von ca. 8o msec, also im Bereich der Merkmalsintegration, müßte die Parallelität auf jeden Fall aufgehoben sein. Daß sich trotzdem der Übungsfortschritt auch auf SOA-Bereiche unterhalb von 8o msec erstreckt, erklären Neisser & Hirst (1974) mit der Bemerkung:

"It seems plausible to suppose that different listeners might c o n s t r u c t different types of analogs in such a task, and also that the nature of the analogs might c h a n g e w i t h p r a c t i c e" (Neisser & Hirst, 1974, S. 394 , meine Hervorhebung).

Wenn aber die analoge Repräsentation selbst wieder eine Konstruktion ist, dann muß 1. geklärt werden, welche Informationen die Grundlage hierfür bilden und 2. durch welche Kodierungsprozesse diese Konstruktion geleistet wird. Zu diesem Punkt fehlen leider explizite Aussagen der Autoren, Die naheliegende Antwort auf die beiden Fragen besteht darin anzunehmen, daß die Informationsbasis hierfür der echoische Speicher ist (Neisser, 1967) - oder Crowder & Morton's (1969) "precategorial acoustic store" bzw. Massaro's (197o) "preperceptual auditory store" -, aber offen bleibt die Frage nach den Kodierungsprozessen, die der Konstruktion der analogen Repräsentation zugrundeliegen. Wenn es richtig ist anzunehmen, daß "the nature of the analogs might change with practice" (Neisser & Hirst, 1974, S. 394), dann müssen offenbar selektive Entnahmeprozesse am echoischen Speicher stattfinden, über deren funktionale Charakteristika die Autoren keine Angaben machen bzw. machen können. Es müßte z.B. geklärt werden, nach welchen Kriterien welche Information selegiert wird und nach welchen Prinzipien sich die Selektion mit Übung bzw. bei sehr kurzen gegenüber längeren SOAs ändert. Das Modell von Neisser

& Hirst (1974) hat einen Teil der Mustererkennungsphänomene nur dadurch "erklären" können, daß ein weiterer erklärungsbedürftiger Verarbeitungsprozeß eingeführt wurde.

Warren (1974b) geht ebenfalls von einem Zwei-Stufen-Modell der auditiven Mustererkennung aus. Wenn die zweite Stufe im Gegensatz zur ersten eine elementweise Kodierung vornimmt, dann sollte eine Erhöhung des SOAs über den Wert hinaus, von dem ab eine elementweise Kodierung bereits möglich ist, die Identifikationsleistung nicht weiter verbessern, zumindest aber nicht verschlechtern. Aber genau letzteres Ergebnis stellte sich ein: Im SOA-Bereich zwischen 127 und 215 msec nahm die Zahl richtiger Identifikationen von Mustern mit größer werdendem SOA monoton zu, sank aber zwischen 215 und 600 msec wieder auf das bei einem SOA von 127 msec erzielte Niveau ab.

Warren (1974b) erklärt den umgekehrt U-förmigen Verlauf der Kurve für die Identifikationsleistung in Abhängigkeit vom SOA durch einen Mustererkennungsprozeß, der die einzelnen Elemente des Reizmusters auf eine Zeitmusterschablone abbildet. Die Rastergröße solcher Schablonen müßte demnach ca. 215 msec betragen, und Abweichungen von dieser Zeitkonstante müssen dann zu unpräzisen Repräsentationen des Gesamtmusters führen.

In einer anderen Arbeit (Warren, 1974a) rückt der Autor wieder von seinem Schablonenkonzept ab und spricht stattdessen von zeitlichen Verbindungen ("temporal compounds") zwischen den Elementen eines Musters, ohne allerdings etwas über die Prozesse der Herstellung solcher Verbindungen auszusagen. Unter zeitlichen Verbindungen versteht Warren solche mit überadditiven Eigenschaften, d.h. die Ver-

bindung der Elemente eines Musters besitzt Eigenschaften, die sich nicht aus der Summe der Eigenschaften seiner Elemente ergeben. Derartig verbundene Elemente können zwar als Verband von anderen Verbänden differenziert werden, aber nicht ohne zusätzliche kognitive Prozesse in ihre Elemente zerlegt werden.

Ähnlich wie Neisser & Hirst (1974) erklärt Warren (1974a) die Unterschiede in der Mustererkennung bei kurzen und langen SOAs durch zwei verschiedene Kodierungsprozesse, die auf der Basis unterschiedlicher Repräsentationen operieren: Eine direkte Identifikation durch elementweise Kodierung der zeitlichen Beziehungen zwischen den Elementen eines Musters ist nur oberhalb von ca. 2oo msec SOA wegen des hohen Zeitbedarfs der Einzelkodierungen möglich; unterhalb von 2oo msec SOA basiert die Mustererkennung nicht mehr auf einzelnen Elementen, sondern auf den Elementverbindungen mit ihren übersummativen Merkmalen, die - ein entsprechendes Lernen vorausgesetzt - den verschiedenen Permutationen der Elemente eines Musters zugeordnet werden können.

"Temporal compounds" und die Neisser'schen "analogs" sind vergleichbare Repräsentationsformen; jedoch besteht ein Unterschied zwischen ihnen in der Art und Weise, wie sie zur Mustererkennung beitragen: Die "analogs" stellen eine Repräsentationsform dar, auf die unabhängig vom SOA eine Stringrepräsentation aufbaut, die dann die Basis für die Reaktionsprogrammierung ist; in Warren's Modell werden die "temporal compounds" in Abhängigkeit vom SOA gebildet, fallen also bei großen SOAs als Zwischenform der Repräsentation fort, während die direkte Identifikation bei Bildung von "temporal compounds" ausfällt. Das Modell von Neisser & Hirst (1974) ist also als l i n e a r e s P r o z e ß m o d e l l der Mustererkennung

anzusehen, während Warren's (1974a) Modell die voneinander unterschiedenen Prozesse als alternative Verarbeitungsstrategien auffaßt, die zu nicht aufeinander aufbauenden Repräsentationen führen.

Auf der Suche nach Begrenzungen der auditiven Mustererkennung hinsichtlich des Erkennens der Reihenfolge der Elemente eines Musters hat sich also herausgestellt, daß eine absolute Begrenzung nicht vorliegt, vielmehr erwiesen sich die zwischenzeitlich als Grenzwerte betrachteten minimalen SOAs immer wieder als relative Grenzwerte, die von den Anforderungen der Aufgabe und der individuellen Erfahrung der Hörer mit der Aufgabe abhängen sowie von der Art der der Mustererkennung zugrundeliegenden Repräsentation der im akustischen Muster enthaltenen Information. Die hierzu formulierten Modelle der Informationsverarbeitung können zwar die Phänomene im großen und ganzen erklären, klammern aber entweder die Prozeßreihenfolge aus, wie im Modell von Warren (1974a), oder können die Prozeßreihenfolge nur dadurch einbeziehen, daß neue erklärungsbedürftige Prozeßannahmen eingeführt werden, wie im Modell von Neisser & Hirst (1974) die Konstruktion der "analogs".

Bis jetzt mußte offen bleiben, ob die Modellvorstellungen auch auf die Wahrnehmung von Musik übertragbar sind, weil die akustischen Muster in den referierten Untersuchungen überwiegend aus zusammenhanglosen Klangereignissen bestanden, denen eine klar definierte Tonhöhe und damit ein wesentliches musikalisches Merkmal fehlt.

Erkennen tonaler Muster in Abhängigkeit von der Größe und Konfiguration der Intervalle

Die über die zeitliche Abfolge hinaus Zusammenhang stiftenden Merkmale von Tonfolgen sind die musikalischen Intervalle und ihre Konfigurationen über die Zeit. Hierbei handelt es sich um diejenigen Merkmale, die am ehesten eine Differenzierung von Geräusch und Musik erlauben, jedoch ihrer zentralen Bedeutung für die Musikwahrnehmung entspricht nicht ihre Berücksichtigung in der Literatur zur auditiven Mustererkennung (Dewar, Cuddy & Mewhort, 1977).

Deutsch (1969) hat als erste ein Modell der musikalischen Mustererkennung vorgestellt und durch zahlreiche Untersuchungen, insbesondere zu interaktiven Effekten zwischen einzelnen Tönen im auditiven Kurzzeitspeicher, zu überprüfen versucht (einen ausführlichen Überblick über diese Arbeiten gibt Deutsch, 1975a, b).

Deutsch's (1969; 198oa) Modell bedient sich einer Analogie zu den neuralen Netzwerken auf dem Niveau rezeptiver Felder des auditiven Cortex und bezieht sich dabei auf die Forschungsergebnisse zu rezeptiven Feldern des visuellen Cortex von Hubel & Wiesel (1962).

Im Modell werden zwei Kanäle angenommen, von denen einer aus sukzessiven Tonfolgen Intervallkodierungen und der andere bei simultanen Tönen Harmoniekodierungen vornimmt. Wir beschränken uns hier auf eine Skizzierung der Mechanismen der Intervallkodierung. Zwei primäre, frequenzspezifische Neuronen konvergieren auf ein sekundäres Neuron, das einem bestimmten Intervall innerhalb des hörbaren Frequenzspektrums entspricht, und mehrere solcher sekundären Neuronen

konvergieren auf tertiäre Neuronen, wenn die Intervallgröße, die sie repräsentieren, identisch ist, die sie aus verschiedenen Bereichen des Frequenzspektrums abstrahiert haben. Auf diese Weise repräsentiert der Output der tertiären Neuronen ein abstraktes musikalisches Intervall unabhängig von seiner Lage im Frequenzspektrum.

Deutsch's Spekulationen zur Neurophysiologie des auditiven Cortex erklären unserer Meinung nach bereits zu viel. Es ist zwar belegt, daß im auditiven Cortex Neuronen vorhanden sind, die auf Frequenzänderungen mit proportionalen Aktionspotentialen antworten und spezifisch für ab- oder aufsteigende Änderungen sind (Suga, 1965; Whitfield & Evans, 1965), aber aus diesen Befunden ist nicht mehr und nicht weniger zu schließen, als daß wir nicht nur konstante Tonhöhen, sondern auch deren Veränderungen in der Zeit wahrnehmen können. Das impliziert aber keineswegs eine Projektion dieser Veränderungen auf das System der musikalischen Intervalle, denn sonst wäre jeder Hörer, ob musikalisch oder nicht, gleich gut in der Lage, Intervalle zu erkennen. Musikalische Intervalle sind unscharfe Kategorien (Burns & Ward, 1973; Siegel & Siegel, 1977a, b), die ein Hörer auf wahrgenommene Frequenzänderung projizieren kann, nicht aber in jedem Moment des Musikhörens auch projizieren muß.

Wenn bereits auf dem Niveau des auditiven Cortex exakte Kodierungen musikalischer Intervalle vorgenommen würden, müßte die Wahrnehmung aller Intervalle unabhängig von ihrer Größe und Position im Frequenzspektrum gleich genau und gleich schwierig sein; die Befunde von Attneave & Olson (1971) sowie Ortmann (1933) sprechen eindeutig dagegen. Außerdem kann dieses Modell die im vorausgegangenen Abschnitt

referierten Befunde zur Abhängigkeit des Erkennens vom SOA nicht erklären, denn wenn Intervallabstraktionen in der beschriebenen Weise stattfinden, dürften für den Hörer bei sehr kleinen SOAs keine anderen Kodierungen notwendig sein, als bei SOAs oberhalb von 2oo msec, bei denen eine elementweise Kodierung möglich zu sein scheint.

Das Modell von Deutsch (1969) berücksichtigt lediglich die Kodierung sukzessiver Töne zu einem Intervall, nicht aber Beziehungskodierungen höherer Ordnung, wie sie offenbar von musikalisch geschulten Hörern vorgenommen werden (Cuddy & Cohen, 1976).

Cuddy & Cohen's (1976) Untersuchung war der Überprüfung des Modells von Deutsch (1969) gewidmet. Deutsch's Modell der Kodierung sukzessiver Intervalle wurden Alternativmodelle gegenübergestellt, die in unterschiedlichem Grade von Intervallinformationen höherer Ordnung zwischen nicht unmittelbar aufeinander folgenden Tönen Gebrauch machen. Das interessanteste Ergebnis dieser Untersuchung bestand in der Demonstration, daß keines der überprüften Modelle den Daten voll gerecht wurde, wobei das Modell von Deutsch mit zu denen gehörte, die die Daten noch am wenigsten erklären konnten. Die Autorinnen vermuten, daß es weniger Intervalle - gleich welcher Ordnung - sind, die beim Erkennen von Melodien verarbeitet werden, sondern daß es sich vielmehr um abstrakte Konfigurationen handeln müsse. Welcher Art diese Konfigurationen sein könnten und wie sie sich beschreiben ließen, haben Cuddy & Cohen (1976) leider nicht ausgeführt. Offenbar meinen sie aber melodische Verlaufsstrukturen, in die Intervalle als Konstituenten eingehen, selbst aber nicht direkt Gegenstand der Wahrnehmung sind.

Einige wenige Untersuchungen sind der Frage der differentiellen Schwierigkeiten der Identifikation tonaler Muster in Abhängigkeit von der durch die einzelnen Töne gebildeten Konfiguration nachgegangen. Die Befunde sind überwiegend so einheitlich, daß sich auf der Basis der relativen Fehlerzahlen bzw. der Erkennungsraten eine Taxonomie der Identifizierbarkeit einfacher tonaler Muster erstellen läßt.

Bereits Ortmann (1926, 1933) untersuchte dieses Problem ausführlich im Kontext von Gedächtnisexperimenten. Seine Einsichten wurden später mit anderen experimentellen Methoden im Rahmen von Mustererkennungsexperimenten aufgegriffen und seine Befunde vielfach repliziert (Divenyi & Hirsh, 1973, 1974, 1975; Kerivan & Carey, 1976; Warren & Byrnes, 1975).

Die Merkmale, die das Erkennen einfacher tonaler Muster zu beeinflussen scheinen, lassen sich aufgrund der Befunde dieser Arbeiten zwei Gesichtspunkten unterordnen: 1. Einflüsse elementarer Merkmale, d.h. solcher Merkmale, die einzelnen Elementen (Tönen und Paaren konsekutiver Töne) zuzuordnen sind, und 2. Einflüsse konfigurativer Merkmale, d.h. solcher Merkmale, die nur das durch die Elemente gebildete Gesamtmuster, nicht aber seine Elemente besitzen.

1. Elementare Merkmale:

a) Anfangs- und Endtöne eines Musters können, wenn sich ihre Tonhöhen deutlich genug von denen anderer Töne abheben, eine Diskrimination zwischen Permutationen derselben Elemente erleichtern (Divenyi & Hirsh, 1974, 1975; Kerivan & Carey, 1976; Ortmann, 1926; Warren, 1972).

b) Intervalle mittlerer Größe (ca. 1/3 bis 2/3 einer Oktave, vgl. Divenyi & Hirsh, 1974) erleichtern das Erkennen eines Musters bzw. seine Diskrimination von anderen Mustern im Vergleich zu deutlich kleineren oder größeren Intervallen (Divenyi & Hirsh, 1974; van Noorden, 1975; Ortmann, 1933).

2. Konfigurative Merkmale:

a) Muster mit Intervallen, die alle in einer Richtung verlaufen (auf- oder absteigend), sind am leichtesten zu erkennen (Divenyi & Hirsh, 1973, 1974, 1975; Kerivan & Carey, 1976; Ortmann, 1933; Warren & Byrnes, 1975).

b) Je mehr Intervallrichtungswechsel ein Muster enthält, desto schwieriger ist das Erkennen (Ortmann, 1933).
Diese Verallgemeinerung Ortmann's wird durch die Befunde von Divenyi & Hirsh (1974) nur partiell gestützt: Da die Autoren nur drei Töne verwandten, konnte maximal auch nur eine Richtungsänderung vorkommen; aber bei Richtungsänderung war die Schwierigkeit des Erkennens deutlich höher als bei gleichgerichteten Intervallen. Die Befunde von Kerivan & Carey (1976) aufgrund von Tonfolgen mit vier unterschiedlich hohen Tönen zeigten ebenfalls eine leichtere Erkennbarkeit der Muster mit gleichgerichteten Intervallen gegenüber Mustern mit Richtungsänderungen. Bei Vergleich derjenigen Muster, die eine oder zwei Richtungsänderungen enthalten, sind die Ergebnisse etwas uneinheitlicher, aber es zeichnet sich eine Tendenz ab, die Ortmann's Verallgemeinerung unterstützt.

Dieses einigermaßen homogene Befundmuster sollte eigentlich geeignet sein, ein Modell der Informationsverarbeitung von Tonfolgen zu formulieren, das nicht nur, wie bisher geschehen, auf der Wahrnehmung der zeitlichen Relationen zwischen Elementen eines Musters basiert (wie bei Divenyi & Hirsh (1974, 1975), Neisser & Hirst (1974) sowie Warren (1974a) und Warren & Byrnes (1975)), sondern darüber hinaus auch die Intervallrelationen und Konfigurationen in einem tonalen Muster berücksichtigt. Die Formulierung und empirische Überprüfung eines solchen Modells steht bislang noch aus.

Die meisten Autoren, die sich mit dem Erkennen von Mustern in Tonfolgen beschäftigt haben, kommen zum Schluß ihrer Arbeiten immer wieder zu der Feststellung, daß bei der Wahrnehmung von Tonfolgen - auch solcher, die noch nicht als Melodien anzusprechen sind - offenbar Gestaltprinzipien eine bedeutende Rolle spielen (z.B. Bregman, 1978a; Divenyi & Hirsh, 1974; Royer & Garner, 1966, 197o). Diese Bemerkung ist dann meistens als Eingeständnis der Unzulänglichkeiten der gegenwärtigen Modellkonstruktion zu verstehen und konstatiert damit nur die Tatsache, daß die Komplexität der Prozesse der auditiven Mustererkennung weitaus höher ist, als ihre Abbildung in Informationsverarbeitungsmodellen bislang erkennen läßt.

2.2.2 ERKENNEN DER MELODIEKONTUR

Der vorausgegangene Abschnitt hat deutlich werden lassen, welche zentrale Funktion das Erkennen von Merkmalen höherer Ordnung für die auditive Mustererkennung besitzt. Wenn das, was Kerivan & Carey (1976) als Konfigurationen und Divenyi & Hirsh (1974)

als Gestalten bezeichnen, durch konfigurative Merkmale ausgezeichnet ist, die u n a b h ä n g i g von elementaren Merkmalen der die Konfiguration oder die Gestalt einer Melodie konstituierenden Elemente wahrnehmbar sind, dann sollte sich diese Unabhängigkeit in Experimenten demonstrieren lassen, in denen Melodien unter verschiedenen musikalischen Transformationen wiederzuerkennen sind. Solche Transformationen können nämlich entweder nur elementare Merkmale betreffen und konfigurative erhalten oder gleichzeitig beide Merkmalstypen tangieren.

Ein Objekt, das aus unterschiedlichen Entfernungen betrachtet wird, behält seine Form unabhängig von der Entfernung des Betrachters. Werner (1926, 194o) untersuchte die auditive Analogie zur Formkonstanz unter der Fragestellung, ob das Charakteristikum optischer Konfigurationen, bei proportionaler Veränderung dieselbe Gestalt beizubehalten, ebenso für tonale Konfigurationen zutrifft, wenn sich ausschließlich die Intervallgrößen ändern. Er konnte in ausführlichen Trainingsversuchen demonstrieren, daß eine Mikrotonskala, in der die Oktave in 72 gleich große Mikrointervalle unterteilt wurde, all die Qualitäten erhält, die die "normale" Tonleiter besitzt (z.B. 1/72-Intervall besitzt Halbtoncharakteristik; die 12. Tonstufe der Mikrotonskala besitzt Leittonfunktion; der 1. und 13. Ton bilden eine Mikro-Oktave usw.). Diese Befunde sprechen erstens gegen den Erklärungswert einer Modellvorstellung der Intervallkodierung, wie sie Deutsch (1969) in ihrem neuralen Netzwerkmodell entwickelt hat, und zweitens für die relativ zur Kodierung von Intervallgrößen wichtigere Funktion von Merkmalen der Konfiguration für die kognitiven Prozesse beim Musikhören.

Ein solches konfiguratives Merkmal ist die K o n t u r einer Tonfolge: das Muster, das unabhängig von der Größe der Intervalle durch den Wechsel des Auf- und Absteigens der Intervalle gebildet wird. In einigen Untersuchungen zum Wiedererkennen von Melodien wurden populäre Lieder oder kurze Tonfolgen verschiedenen Formen von Transformationen unterzogen und die Wiedererkennungswahrscheinlichkeiten in Abhängigkeit der sich durch die Transformation ändernden Merkmale einer Melodie erhoben. Ziel dieser Untersuchungen war es, diejenigen Informationen in Tonfolgen zu identifizieren, die im Gedächtnis längerfristig gespeichert sein könnten. Da es sich herausgestellt hatte, daß Informationen über absolute Tonhöhen bereits nach wenigen Sekunden nicht mehr verfügbar sind (z.B. Darwin, Turvey & Crowder, 1972; Deutsch, 197o; Siegel, 1972; Wickelgren, 1966, 1969), mußte davon ausgegangen werden, daß im Kurzzeit- und Langzeitgedächtnis nur Informationen über die Ergebnisse der Weiterverarbeitung der Tonhöheninformationen gespeichert sind.

Die weitgehend übereinstimmenden Befunde der Einzelarbeiten lassen sich folgendermaßen zusammenfassen (Deutsch, 1972; Dowling, 1971, 1972, 1978; Dowling & Fujitani, 1971; Dowling & Hollombe, 1977; House, 1977; Idson & Massaro, 1978; Kallman & Massaro, 1979; White, 196o):

1. Konturerhaltende Transformationen beeinträchtigen das Wiedererkennen einer bekannten Melodie weitaus weniger als konturzerstörende Transformationen.

2. Bei konturerhaltenden Transformationen hängt die Schwierigkeit des Wiedererkennens vom Ausmaß der Veränderung elemen-

tarer Merkmale ab: Je weniger elementare Merkmale verändert werden, desto leichter kann eine Melodie wiedererkannt werden.

a) Die geringsten Schwierigkeiten machen Transpositionen, da hierbei bis auf die individuellen Tonhöhen alle weiteren Merkmale erhalten bleiben.

b) Bei Verschiebung entlang der diatonischen Tonleiter (z.B. die Tonfolge c e g h, bestehend aus der aufsteigenden Intervallfolge große Terz, kleine Terz und große Terz, wird durch Verschiebung um eine Terz aufwärts bei Beibehaltung der Tonart zur Tonfolge e g h d, die nun aber aus der Intervallfolge kleine Terz, große Terz, kleine Terz besteht) entstehen zusätzlich Veränderungen in der Größe der aufeinanderfolgenden Intervalle; in diesem Fall ist ein Wiedererkennen einer Melodie schwieriger als bei Transposition.

c) Bei rein zufälliger Änderung der Intervallgrößen unter Beibehaltung der Intervallrichtungen ist ein Wiedererkennen bekannter Melodien noch schwieriger als bei Verschiebung.

3. Bei konturzerstörenden Transformationen hängt die Schwierigkeit des Wiedererkennens sowohl von der Art als auch von der Anzahl der Transformationen ab.

a) Die relativ geringsten Schwierigkeiten bereitet die Umkehrungstransformation, durch die - unter Beibehaltung der exakten Intervallgrößen - die Intervallrichtungen umgekehrt werden.

b) Schwieriger als bei Umkehrungstransformation ist das Wiedererkennen bei Spiegelungstransformation. Hierbei wird die Reihenfolge der Töne in der Zeit umgekehrt.

c) Werden beide Transformationen gleichzeitig angewandt, so ist die Melodie schwieriger wiederzuerkennen, als wenn sie nur eine Transformation enthielte.

Wenn für das Wiedererkennen von Melodien ihre konfigurativen Merkmale eine zentralere Rolle als ihre elementaren Merkmale spielen - Davies & Jennings (1977) gehen sogar so weit zu behaupten, daß n u r ihre konfigurativen Merkmale kritisch sind -, dann sollte es sich hierbei auch um diejenigen Merkmale handeln, die für die Prozesse der auditiven Mustererkennung eine dominante Rolle spielen, denn dem Wiedererkennen muß ja notwendigerweise eine zumindest partielle Identifikation vorausgehen.

Ein funktionales Modell der Prozesse der auditiven Mustererkennung muß daher dieser "Dominanz" konfigurativer Merkmale Rechnung tragen. Dies ist zwar auch im Rahmen der Informationsverarbeitungsmodelle annäherungsweise geschehen, indem man die Wahrnehmung von Konfigurationen durch Abstraktions- oder Konstruktionsprozesse zu erklären versuchte, deren Input elementare Merkmale des Reizmusters darstellen (z.B. Deutsch, 1978), diese Erklärung ist aber insofern unzutreffend, als Cuddy & Cohen (1976), insbesondere aber Davies (1979), Davies & Jennings (1977) sowie Davies & Yelland (1977) überzeugende Argumente und experimentelle Belege dafür ins Feld führen konnten, daß offenbar elementare Merkmale, wenn sie zur Aufgabenbewältigung nicht benötigt werden, auch gar nicht repräsentiert werden, ohne dadurch die Wiedererkennungsleistung oder die Reproduktion zu beein-

trächtigen. Offenbar können konfigurative Merkmale von Melodien ohne den Umweg der Konstruktion aus elementaren Merkmalen d i r e k t wahrgenommen und kodiert werden; dies impliziert aber nicht notwendigerweise, daß die Wahrnehmung und Kodierung elementarer Merkmale in allen Wahrnehmungssituationen von untergeordneter oder gar keiner Bedeutung ist, wie Davies & Jennings (1977) annehmen.

Eine Erklärung dieser Phänomene im Rahmen eines linearen Informationsverarbeitungsmodells stößt gerade wegen der Linearitätsannahme über die Prozeßabfolge auf einige Schwierigkeiten. Wir werden auf diesen Punkt noch einmal zurückkommen, wenn wir einen alternativen Erklärungsansatz vorstellen, der auf die Linearitätsannahme verzichten kann.

2.2.3 ORGANISATION BEI WAHRNEHMUNG UND SPEICHERUNG VON TONFOLGEN

Wir haben in den beiden vorangegangenen Abschnitten die auditive Mustererkennung unter dem Aspekt der Merkmale des Reizmusters betrachtet, auf die sich in unterschiedlichen Aufgabensituationen (Erkennen der Reihenfolge bzw. Diskrimination einer Tonfolge von anderen; Wiedererkennen einer Tonfolge) die Entscheidungen der Vpn stützen. Mit der Identifikation von Merkmalen kann die Mustererkennung allerdings in solchen Aufgabensituationen nicht abgeschlossen sein, in denen längere Abschnitte einer Tonfolge miteinander in Beziehung gesetzt werden müssen oder eine Tonfolge längerfristig gespeichert werden muß. In diesen Fällen ist eine Integration der durch die Merkmalsanalyse gelieferten "Bruchstücke" zu einer organisierten Gesamtrepräsentation der Tonfolge notwendig. Welche kognitiven

Prozesse bei der Herstellung einer Repräsentation beteiligt sind und welche Organisationsstrukturen solche Repräsentationen besitzen können, wollen wir im folgenden Abschnitt darstellen. Hierbei müssen wir die Frage nach der gemeinsamen Klammer beantworten, die einer Folge isolierter Töne den phänomenalen Status eines sich in der Zeit entfaltenden Musters, einer Konfiguration oder einer Gestalt verleiht.

Da es sich hierbei um Integrationsleistungen kognitiver Systeme handelt, die sich über die Zeit erstrecken, firmiert ein Teil der zu berichtenden Modelle und Befunde in der Literatur häufig unter dem etwas irreführenden Titel "Zeitwahrnehmung". Irreführend ist diese Bezeichnung des Forschungsbereichs deshalb, weil die Zeit nicht wie ein Reiz wahrnehmbar ist, sondern überhaupt erst durch sich in der Zeit abspielende Ereignisse ins Bewußtsein tritt: Zeitwahrnehmung ist die Wahrnehmung von Veränderungen in der Zeit (Fraisse, 1978). Doch ist es weniger der Aspekt der "Wahrnehmung" von Zeit, der uns im gegenwärtigen Kontext interessiert, als vielmehr der komplementäre Aspekt des durch die Wahrnehmung in einer zeitlichen Abfolge gestifteten Zusammenhangs von isolierten, zeitlich voneinander getrennten Reizen. In Untersuchungen zur Zeitwahrnehmung steht die Bestimmung der Größe von Zeitintervallen im Vordergrund, innerhalb derer gerade noch oder eben erst eine bestimmte Integrationsleistung durch kognitive Systeme vollzogen werden kann; wir konzentrieren uns dagegen auf Aussagen über diese Integrationsleistungen selbst, um etwas über die Prozesse, die diesen Systemen zugeschrieben werden, zu erfahren.

Integration und Segmentierung

Integrative Prozesse finden auf allen Stufen der Informationsverarbeitung statt; wir haben bereits die Energieintegration und die Merk-

malsintegration im Zusammenhang mit der auditiven Mustererkennung erwähnt. Hierbei handelt es sich um integrative Leistungen kognitiver Systeme, die auf Verarbeitungsstufen vor oder während der echoischen Speicherung anzusiedeln sind (Massaro, 1975). Die in der Verarbeitungssequenz nachfolgenden Integrationen unterscheiden sich von denen der vorausgehenden sowohl durch die größere Zeitspanne, über die sich die Integration erstreckt, als auch durch die Funktionsweise der daran beteiligten kognitiven Prozesse.

Zeitliche Integrationen werden in der Literatur zur Zeitwahrnehmung zwei unterschiedlichen phänomenalen Modi des Zeiterlebens zugeschrieben: Ornstein (1969) unterscheidet u.a. zwischen der psychischen Präsenzzeit, "in which we experience the flow of events as being simultaneously available to perceptual or cognitive analysis" (Michon, 1978, S. 9o), und der erlebten Dauer, die aufgrund einer Rekonstruktion zeitlich getrennter Gedächtnisinhalte erschlossen wird. Toda (1975) ordnet diesen Modi des Zeiterlebens zwei kognitive Subsysteme zu: Das primäre System ist ständig aktiv, arbeitet auf der Grundlage sukzessiver Wahrnehmungen und benötigt keinen Zugriff auf Inhalte des Kurz- oder Langzeitgedächtnisses, während das sekundäre System nicht ständig aktiv ist und auf der Grundlage von Gedächtnisinhalten zeitliche Informationen erschließt.

Michon (1978) definiert die psychische Präsenzzeit im Sinne von James (189o) als diejenige Zeitspanne, innerhalb der die Konstituenten eines sukzessiven Reizmusters s i m u l t a n v e r f ü g b a r und einer Weiterverarbeitung zugänglich werden, wenn der Struktur des Reizmusters fokale Aufmerksamkeit zugewandt wird. Die Bestimmungsstücke dieser Definition können wir auch zur Defi-

nition des Begriffs der Integration im hier zu verwendenden Sinne heranziehen: Integration ist diejenige Wahrnehmungsleistung oder Leistung höherer kognitiver Systeme, die sukzessive Wahrnehmungen innerhalb des als psychische Präsenzzeit bezeichneten Zeitraumes simultan verfügbar machen und einer Weiterverarbeitung als Ganzes zuführen, wenn ihrer Struktur fokale Aufmerksamkeit zugewandt wird.

Fraisse (1978) gibt den gleitenden Zeitausschnitt der psychischen Präsenzzeit für längere Reizmuster mit maximal 5 sec. an. In dieser Größenordnung liegt auch die zeitliche Erstreckung eines Taktes innerhalb des langsamsten Adagios.

Der der Integration zugrundeliegende Funktionsmechanismus wird als Einheitenbildung im Sinne des Chunking (Miller, 1956a, b) beschrieben (Fraisse, 1978; Michon, 1978). Michon (1978) verwendet allerdings überwiegend den Begriff der Segmentierung. Wir verwenden das Begriffspaar "Integration/Segmentierung" deskriptiv zur Bezeichnung zweier komplementärer Aspekte des kognitiven Funktionsmechanismus des Chunkings: Der Terminus "Integration" bezeichnet den Aspekt des S t i f t e n s von Beziehungen zwischen den Elementen eines Reizmusters, während der der "Segmentierung" den Aspekt der A b g r e n z u n g der durch einen Chunkingprozeß integrierten Elemente beschreibt.

In diesem Sinne verwenden wir auch den Begriff der Integration von Wahrnehmungen zur Beschreibung integrativer Leistungen, die sich unmittelbar auf Wahrgenommenes beziehen, z.B. zur Beschreibung des Erlebens der Zusammengehörigkeit von vier Tönen eines 4/4-Taktes; hierbei handelt es sich also um Integrationen, die Toda's (1975) primärem System zugeschrieben werden können. Integrationen können sich aber

auch auf Gedächtnisinhalte beziehen, die zeitlich soweit voneinander getrennt gebildet wurden, daß sie nicht mehr in die Zeitspanne der psychischen Präsenzzeit fallen. In diesem Fall sprechen wir deskriptiv von einer Integration von Gedächtnisinhalten, die Toda's (1975) sekundärem System zugeordnet werden kann, wie z.B. beim Herstellen einer Beziehung zwischen der Exposition und der Reprise eines Sonatenkopfsatzes.

Bedingungen für die Integration sukzessiver Wahrnehmungen

Beziehungen zwischen einzelnen Tönen einer Tonfolge oder Melodie werden durch die zeitliche Nähe oder ihre phänomenale Kohärenz (van Noorden, 1975), durch den Rhythmus und seine prosodische Struktur und/oder durch die Ähnlichkeit der Tonhöhen in einer Tonfolge hergestellt.

(1) Zeitliche Nähe und Kohärenz

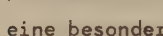
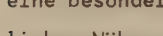
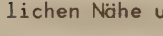
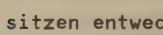
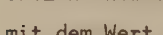
Als das Phänomen der Kohärenz der Töne einer Tonfolge bezeichnet van Noorden (1975) "the connection observed in the perception of a tone sequence, characterized by the fact that the observer has the impression that the tone sequence is a connected series of tone bursts, ordered in time" (van Noorden, 1975, S. 2).

Das Phänomen der Kohärenz soll sich bei SOAs zwischen ca. 35o und 7oo msec besonders deutlich einstellen und sowohl bei kürzeren als auch bei längeren SOAs wieder schwächer erlebt werden (Fraisse, 1956; van Noorden, 1975). Dieser von Michon (1978) angegebene SOA-Bereich stellt

allerdings nur eine sehr grobe Schätzung dar, die bereits durch interaktive Effekte mit anderen Faktoren als der zeitlichen Nähe beeinflußt ist; wenn man das als Interaktionseffekt von SOA und Intervallgröße aufzufassende "auditory streaming" (vgl. Bregman, 1978a) aus dieser Schätzung herausläßt, kann sich das Erlebnis von zeitlicher Kohärenz unabhängig von anderen Faktoren in einem weitaus größeren Zeitintervall einstellen, nämlich innerhalb des von Fraisse (1978) angegebenen Grenzwertes für die zeitliche Ausdehnung der psychischen Präsenzzeit von ca. 5 sec. Innerhalb dieses Zeitintervalls nimmt die erlebte Kohärenz bei Zunahme der zeitlichen Nähe zwischen den Elementen eines Reizmusters zu, wird aber gleichzeitig um so anfälliger für Organisationseffekte aufgrund der Ähnlichkeit der Tonhöhen, wie dies beim "auditory streaming" der Fall ist.

Wenn Kohärenz aufgrund zeitlicher Nähe einen bedeutenden Faktor für integrative Leistungen kognitiver Systeme darstellen soll, dann muß eine Variation der zeitlichen Nähe (operationalisiert durch eine Variation der SOAs zwischen Elementen) die Elemente eines Reizmusters zu Gruppen organisieren, innerhalb derer eine höhere Kohärenz erlebt wird als zwischen ihnen.

Ein systematisches Einfügen von Pausen zwischen einige Töne einer Tonfolge kann die Identifikation der Struktur (hier als Reihenfolge der Töne verstanden) wesentlich erhöhen, wenn die dadurch erzielte Organisation in Tongruppen mit der durch Variation der Tonhöhen erzielten Organisation kompatibel ist; führt die Positionierung der Pausen dagegen zu einer inkompatiblen Organisation, sinkt die Identifikationsrate unter das Niveau, das für die gleiche Tonfolge ohne eingefügte Pausen erzielt wird (Handel, 1973; Handel & Yonder, 1975; McNally & Handel, 1977).

Eine echte Melodie zeichnet sich im Unterschied zu in vielen Experimenten zur auditiven Mustererkennung verwandten Tonfolgen durch eine besonders regelmäßige Variation der SOAs und damit der zeitlichen Nähe und der erlebten Kohärenz aus: Die einzelnen SOAs besitzen entweder eine Größe, die eine Zweierpotenz multipliziert mit dem Wert des kleinsten in der Melodie vorkommenden SOAs beträgt (z.B. gilt diese Beziehung bei Vergleich der Takte  und ), oder die ein ganzzahliges Vielfaches des 1 1/2-fachen des kleinsten SOAs beträgt (z.B. bei Vergleich der Takte  und  oder bei einfachen Synkopierungen, wie z.B. innerhalb des Taktes ). Diese Regelmäßigkeit in der Variation der zeitlichen Nähe von Tönen in Melodien stellt damit eine ausgezeichnete Basis für Integration bei der Wahrnehmung von Tönen dar, zwischen denen die SOAs gleich groß sind, und für eine Segmentierung in Tongruppen, zwischen denen das SOA größer als innerhalb einer Tongruppe ist. Diese Form der Regelmäßigkeit ist eine notwendige Bedingung für die Wahrnehmung von Rhythmen in Musik.

(2) Rhythmus und prosodische Struktur

Um die Wahrnehmung von Rhythmus zu gewährleisten, muß eine Tonfolge eine zeitlich regelmäßige Akzentstruktur (prosodische Struktur) besitzen, die unabhängig vom Tempo der Tonfolge konstant bleibt (Martin, 1972). Voraussetzung für die Wahrnehmung von Rhythmen ist ein Mindestmaß an zeitlicher Nähe zwischen den Tönen einer Melodie, nicht aber unbedingt eine besonders ausgezeichnete Regelmäßigkeit der zeitlichen Nähe, obwohl diese in Musik immer existiert.

Wenn die in einer Tonfolge enthaltenden Akzente (größere Lautstärke akzentuierter gegenüber nicht akzentuierten Tönen) jeweils auf den ersten Ton einer integrierten Gruppe von Tönen fallen, dann ist die Identifikation der Struktur einer Tonfolge wesentlich einfacher, als vergleichsweise für Tonfolgen, in denen die Akzente an anderer Stelle innerhalb einer Gruppe auftreten (Handel & Yonder, 1975; Sturges & Martin, 1974). In letzterem Fall entstehen dann nämlich zwei alternative Möglichkeiten der Segmentierung, die in Konflikt geraten können: Entweder können Töne aufgrund der zeitlichen Nähe zu einer Gruppe integriert werden oder es werden Gruppen organisiert, deren Größe und Abgrenzung durch die Akzentfolge bestimmt ist. In diesem Konfliktfall setzt sich dann allerdings die Organisation aufgrund der Variation der zeitlichen Nähe immer durch (Handel & Yonder, 1975).

An dieser Stelle wird die Notwendigkeit einer Differenzierung zwischen Integration und Segmentierung deutlich: Das Durchsetzen der Organisation auf der Basis zeitlicher Nähe gegenüber einer möglichen Organisation auf der Basis deplazierter Akzente, ist dadurch zu erklären, daß zeitliche Nähe für die kognitive Verarbeitung eine Bedingung für zeitliche Integration darstellt, während Akzente ausschließlich eine Bedingung für Segmentierung zeitlich integrierter Töne darstellen. Durch deplazierte Akzente werden die durch zeitliche Nähe gebildeten Gruppen nicht aufgehoben, sondern erhalten lediglich eine gruppeninterne Struktur, die Teilgruppen segmentiert.

Handel & Yonder (1975) benutzten in ihren Tonfolgen lediglich zwei verschiedene Frequenzen, ließen die sich beim Hören einstellenden

Organisationen beschreiben und schlossen hieraus auf die Art vorgenommener Segmentierungen. Eine direktere Methode zum Nachweis rhythmischer Gruppenbildungen wandte Dowling (1973a) bei relativ komplexen Tonfolgen an. Er bot den Vpn rhythmisch organisierte Tonfolgen, die als vier Gruppen von je fünf Tönen organisiert werden konnten. Zwei Sekunden später wurden den Vpn fünf weitere Töne vorgespielt, und sie mußten entscheiden, ob es sich hierbei um einen Ausschnitt aus der dargebotenen Tonfolge handelt oder nicht. Die relative Zahl richtiger Entscheidungen war bei Darbietung einer vollständigen Gruppe höher als bei Darbietung der letzten drei Töne einer Gruppe und der ersten beiden der unmittelbar folgenden Gruppe.

Die von Dowling (1973a) verwandten Tonfolgen bildeten einen 6/8 oder 3/4 Takt, wobei eine Gruppe immer genau einem Takt entsprach. Da sich die prosodische Struktur in Musik immer taktweise wiederholt, liegt die Hypothese nahe, daß Takte funktionale Verarbeitungseinheiten darstellen.

Eine weitere für die Mustererkennung in Tonfolgen kritische Bedingung rhythmischer Organisation ist die relative Stärke einzelner Akzente. Wird die Akzentuierung einzelner Töne durch eine zu starke Anhebung der Lautstärke gegenüber nicht akzentuierten Tönen produziert, dann tritt bei gleichzeitig relativ hohen Darbietungsgeschwindigkeiten "perceptual streaming" (oder wie van Noorden (1975, 1977) es nennt: "perceptual fission") auf: Akzentuierte und nicht akzentuierte Töne (auch bei konstanter Frequenz) bilden getrennte Tonfolgen, die in einem Figur-Grund-Verhältnis zueinander stehen (van Noorden, 1975, 1977; Pinheiro & Ptacek, 1971; Ptacek & Pinheiro, 1971).

Dieses Phänomen der Aufspaltung einer Tonfolge in getrennte Teiltonfolgen war die Ursache für die besonders große Schwierigkeit bei der Identifikation von Tonfolgen in einer der experimentellen Bedingungen der Arbeit von Handel & Yonder (1975): Bei deplazierten Akzenten und gleichzeitig hoher Darbietungsgeschwindigkeit ergab sich eine deutlich geringere Identifikationsrate als bei geringer Darbietungsgeschwindigkeit. Interessant ist, daß dieser Effekt nicht auch bei hoher Darbietungsgeschwindigkeit, aber richtig plazierten Akzenten auftrat. Offenbar ist die Kohärenz der durch zeitliche Nähe integrierten Töne so ausgeprägt, daß sich eine alternative Organisation nicht durchzusetzen vermag. Das, was Werner (1919) als "Gestaltverkettung" im Zusammenhang mit der Wahrnehmung rhythmischer Organisationen bezeichnet hat, nämlich die Zusammenhangbildung zwischen rhythmischen Einheiten, wird offenbar nur durch richtig plazierte Akzente unterstützt, die auf den ersten Ton einer durch zeitliche Nähe integrierten Gruppe fallen.

Eine Folge identischer Töne, die mit konstantem SOA und ohne Akzente aufeinander folgen, können auch als rhythmisch organisiert erlebt werden. Wundt (1874) spricht hierbei von subjektiver Rhythmisierung. Bolton (1894) untersuchte die Bedingungen für ihr Auftreten und stellte fest, daß ein subjektiver Rhythmus nur dann erlebt wird, wenn die einzelnen Reize als distinkt (bei SOAs größer oder gleich 115 msec), aber nicht schon als unabhängig wahrgenommen werden (was bei SOAs größer oder gleich 158o msec nach Bolton (1894), zwischen 1.5 und 2 sec nach McDougall (19o3) und Fraisse (1956) der Fall wäre (SOA-Angaben nach Fraisse, 1978)). Das offenbar optimale SOA für subjektive Rhythmisierung liegt bei ca. 4oo msec (Meumann, 1894; Fraisse, 1956).

Subjektiv rhythmisierte Tonfolgen werden wie durch Akzentvariationen und/oder SOA-Variationen rhythmisierte erlebt: Die Zeitintervalle zwischen subjektiven Gruppen erscheinen größer als die Zeitintervalle innerhalb einer Gruppe, und bei längerer Darbietung erscheint allmählich das erste Element einer Gruppe akzentuiert (Bolton, 1894; Koffka, 1909; Miyake, 1902; Stetson, 1905; Wirth, 1937).

Obwohl in akzentuierten Tonfolgen die Grundlage der rhythmischen Segmentierung von Tongruppen reizimmanent ist und dies bei subjektiver Rhythmisierung nicht der Fall ist, legen diese phänomenalen Übereinstimmungen des rhythmischen Erlebens die Hypothese nahe, daß es sich in beiden Fällen um funktional sehr ähnliche Prozesse handeln könnte. Aber im Gegensatz zur Rhythmuswahrnehmung in akzentuierten Tonfolgen stellt die subjektive Rhythmisierung einen a k t i v e n S e g m e n t i e r u n g s p r o z e ß des Hörers dar (Fraisse, 1978).

Bisher haben wir rhythmische Segmentierungen immer nur als eine lineare Folge von akzentuierten Gruppen betrachtet; aber gerade musikalische Rhythmen, die durch Akzente mindestens zweier Abstufungen ausgezeichnet sind (Hauptbetonungen und Nebenbetonungen), lassen sich als hierarchisch geordnet betrachten (Cooper & Meyer, 1960; Martin, 1972). Experimentelle Belege für eine hierarchische Organisation bei der Wahrnehmung rhythmisch segmentierter Tonfolgen können die Untersuchungen von Perkins (1974) und Sturges & Martin (1974) liefern.

(3) Ähnlichkeiten zwischen Tönen aufgrund ihrer Tonhöhe

Auch in Abwesenheit von systematischen Akzentvariationen und einer Variation der zeitlichen Nähe zwischen Tönen werden Tonfolgen als organisiert wahrgenommen, wenn systematische Tonhöhenunterschiede vorhanden sind. Bregman (1978a) und Fraisse (1978) sprechen bei Organisation aufgrund von Tonhöhenunterschieden von einer Organisation nach dem Prinzip der Ähnlichkeit: Töne gleicher oder nur geringfügig unterschiedlicher Höhe bilden eine gemeinsame Gruppe, und einzelne Gruppen werden durch Tonhöhenunterschiede voneinander getrennt, wenn diese größer als die innerhalb einer Gruppe sind.

In Untersuchungen zur Organisation von Tonfolgen nach dem Ähnlichkeitsprinzip wurden zwei verschiedene experimentelle Ansätze verfolgt: 1. Die Beschreibung der wahrgenommenen Struktur binärer Tonfolgen (Garner (1974) und Mitarbeiter; Handel (1974) und Mitarbeiter) und 2. Wahrnehmung tonhöhenabhängiger Organisationen bei schneller Darbietungsgeschwindigkeit (Bregman (1978) und Mitarbeiter).

Hört eine Vp eine ununterbrochen wiederholte Tonfolge, die aus zwei verschieden hohen Tönen besteht, z.B. aus den alternierenden Paaren eines hohen (H) und eines tiefen (T) Tons, dann wird diese Tonfolge entweder als H-H-T-T oder als T-T-H-H, nicht aber als H-T-T-H oder T-H-H-T wahrgenommen. Identische Elemente werden zu einer Gruppe integriert, die Anfang und Ende des wahrgenommenen Musters bilden.

Preusser, Garner & Gottwald (1970a, b) sprechen davon, daß diese Organisation eine Figur-Grund-Differenzierung darstelle, in der

die längere Tonfolge identischer Elemente die Figur und die kürzere den Grund bilde (wenn mehr Töne als in unserem Beispiel verwendet werden). Handel (1974) und Handel & Pickens (1977) stellten fest, daß überwiegend die Gruppe der tiefen Töne den Anfang der Beschreibung einer Vp und auch die Figur bilden, Royer & Garner (197o) fanden dagegen bei ihren Vpn eine Bevorzugung einer Musterbeschreibung, die mit der Gruppe der hohen Töne beginnt, die auch die Figur bildet. Übereinstimmendes Ergebnis dieser Untersuchungen ist der Befund, daß immer solche Gruppen organisiert werden, die zu möglichst langen Folgen identischer Elemente führen, wobei die Abfolge der Länge einzelner Gruppen systematisch zu- oder abnehmend ist. Die Wahrnehmung tendiere, so Royer & Garner (197o), offenbar zu möglichst e i n f a c h e n O r g a n i s a t i o n e n , die sich über die gesamte Länge eines Musters erstrecken; deshalb müsse auch die Wahrnehmung zeitlicher Muster auf einem g a n z h e i t l i c h e n P r o z e ß beruhen, der nicht als sequentielle Verknüpfung einzelner Elemente verstanden werden könne.

Garner (1962) stellte die Hypothese auf, daß bei einer ganzheitlichen Wahrnehmung ein Muster mit einer guten Gestalt weniger Organisationsalternativen besitzen sollte, als ein Muster, das keine gute Gestalt besitzt. Gleichzeitig sollte dann auch die Schwierigkeit des Mustererkennens bei Vorliegen nur weniger Organisationsalternativen geringer als bei Vorliegen vieler sein. Diese Erwartung wurde durch Befunde von Garner & Clement (1963) sowie Royer & Garner (1966) bei Darbietung optischer Reize bestätigt. Später konnte dieser Befund von Royer & Garner (197o) auch mit akustischen Signalen repliziert werden.

Wir können die von Garner und Mitarbeitern sowie von Handel und Mitarbeitern demonstrierten Organisationseffekte ebenfalls als Integrationen bzw. Segmentierungen beschreiben, denn zeitlich nahe, ähnliche Töne bilden eine Gruppe, und die Gruppen werden durch die unterschiedliche Tonhöhe ihrer Elemente an der Grenze zwischen Gruppen getrennt. Ob allerdings diese Organisationseffekte auf die spezifischen experimentellen Bedingungen zurückzuführen sind, die ja ein wiederholtes Hören der Tonfolgen gestatteten, oder ob diese Effekte auch bei einmaligem Hören ihre "Ganzheitlichkeit" zeigen, läßt sich mangels einschlägiger Vergleichsuntersuchungen nicht beantworten.

Diese gerade geschilderten Organisationsformen binärer Tonfolgen erwiesen sich auch bei recht unterschiedlichen Darbietungsgeschwindigkeiten als äußerst stabil (Garner & Gottwald, 1967, 1968). Royer & Garner (197o) und Preusser (1972) berichten jedoch, daß bei sehr hohen Darbietungsgeschwindigkeiten eine Figur-Grund-Differenzierung auftritt, die sich phänomenal darin äußert, daß die hohen und die tiefen Töne als eigenständige Muster wahrgenommen werden. Hierbei handelt es sich um das gleiche Phänomen, auf das Miller & Heise (195o; Heise & Miller, 1951) im Rahmen ihrer Untersuchungen zur Trillerschwelle stießen und das später von Bregman & Campbell (1971) als "primary auditory stream segregation effect" bezeichnet wurde.

Musikern ist dieses Phänomen unter dem Namen "implizite Polyphonie" (Bukofzer, 1947) geläufig. Hierbei handelt es sich um eine Kompositionstechnik, die insbesondere in Werken für Soloinstrumente ohne Begleitung in der Zeit des Barock zu finden ist.

Dabei spielt ein einzelnes Instrument scheinbar gleichzeitig zwei Stimmen. Dieser Eindruck entsteht dadurch, daß sich hohe und tiefe Töne in schneller Folge abwechseln. Obwohl bei mehrmaligem Hören der Sukzessivcharakter der Töne, die verschiedenen Stimmen zugeordnet sind, bei einiger Konzentration auf seiten des Hörers deutlich wird, ist er selbst bei kurzen Melodien nicht in der Lage, die Reihenfolge der Töne richtig zu erkennen.

Beim Hören einer derart "auseinandergebrochenen" Melodie erkennt man immer nur eine Melodie auf dem Hintergrund der anderen, nie aber können beide gleichermaßen im Vordergrund stehen. Es ist allerdings bei erneutem Hören möglich, das Verhältnis von Figur und Grund umzukehren (Dowling, 1973b).

Bregmann & Campbell (1971) sahen einen Zusammenhang zwischen den Schwierigkeiten beim Erkennen der Reihenfolge von vier qualitativ unterschiedlichen Geräuschen bei hohen Darbietungsgeschwindigkeiten in der Untersuchung von Warren et al. (1969) und den Befunden von Miller & Heise (195o), Norman (1967) und Peters (1964, 1967), wonach die Reihenfolge von Tönen bei hohen Darbietungsgeschwindigkeiten nicht richtig zu erkennen ist, wenn die Intervalle zwischen den Tönen relativ groß sind.

Streaming zeigt sich als eine Interaktion von Darbietungsgeschwindigkeit und Intervallgröße zwischen sukzessiven und (in schwächerem Maße) weiter entfernten Tönen (Bregman & Dannenbring, 1973; Dannenbring & Bregman, 1976 a; Massaro, 1977): Je größer die Darbietungsgeschwindigkeit, desto geringer kann der Frequenzunterschied zwischen denjenigen Tönen sein, die noch gerade eben als zwei getrennte Tonfolgen wahrgenommen werden, wobei das SOA, nicht das ISI ausschlaggebend ist (Dannenbring & Bregman, 1976b (Ähnliche Befunde ergaben Untersuchungen zur Identifikation von Vokalreihenfolgen, vgl. Dorman, Cutting & Raphael (1975)).

Während Dowling (1973b), Idson & Massaro (1976) und Norman (1967) den Streaming-Effekt auf die begrenzte Kapazität der zentralen Verarbeitungsinstanz zurückführen, sehen Bregman und Mitarbeiter (Bregman, 1978a, b; Bregman & Rudnicki, 1975; Dannenbring & Bregman, 1976a, b) hierin eine p o s i t i v e L e i s t u n g des auditiven Systems, die dazu dient, verschiedene Schallquellen, die gleichzeitig aus derselben Richtung zu kommen scheinen, voneinander zu trennen, so daß sich die Aufmerksamkeit auf eine dieser Quellen selektiv richten kann: "... the aspects that can be brought together in a single act of attention are not at all arbitrarily selectable. We are built to pay attention to auditory sources, not to acoustic components, and it is the decomposition heuristics, organized in ways described by Gestalt psychology, that build the auditory descriptions of sources out of complex acoustic input and, in so doing, place strong constraints on the process of attention" (Bregman, 1978a, S. 74).

Wie ein Vergleich der Organisation von Tonfolgen aufgrund von Ähnlichkeiten zwischen Tönen mit der Streamorganisation zeigt, sind hier analoge Organisationsprinzipien realisiert: ähnliche Töne werden zu einer Gruppe bzw. einem Stream vereinigt, und geringe Ähnlichkeiten zwischen Tongruppen markieren Gruppengrenzen bzw. differenzieren Streams voneinander. Da die "Dekompositionsheuristiken", die Bregman (1978a) für die Bildung von Streams beschreibt, offenbar den Prinzipien der Gruppenorganisation entsprechen, liegt die Vermutung nahe, daß es sich in beiden Fällen auch um dieselben Verarbeitungsmechanismen handeln könnte.

Diese Hypothese ist auch bei funktionaler Betrachtung der beiden Phänomene plausibel. Neumann (198o) sieht eine Parallele zwischen den für die Streamorganisation kritischen Merkmalen und anderen physischen Reizmerkmalen, die als Selektionskriterien für denjenigen Modus der Aufmerksamkeit fungieren, den er "konstantes Fokussieren" nennt. Hierbei handelt es sich für die auditive Modalität um Merkmale der räumlichen Lokalisation, z.B. beim dichotischen Hören (Cherry, 1953; Cherry & Taylor, 1954; Spieth, Curtis & Webster, 1954), Intensität und Frequenzbereich (z.B. Egan, Caterette & Thwing, 1954) und Stimmqualität bei überlagerten Stimmen (z.B. Broadbent, 197o). Neumann (198o) sieht im Streaming genauso eine fehlende Einheit der Tonfolge als Wahrnehmungsgegenstand, wie sie bei unterschiedlich lokalisierten Schallquellen im Hörraum vorliegt; oder andersherum betrachtet handelt es sich hierbei um diejenigen Merkmale, die bei ähnlichen Frequenzen oder gleicher Lokalisation von Signalen im Hörraum O b j e k t p e r m a n e n z stiften (Neumann, 198o) und dadurch einen Gegenstand oder eine Schallquelle fokussierbar machen (Bregman, 1978a).

Die durch Integration und Segmentierung gebildeten Gruppen stellen also Einheiten der auditiven Modalität dar, die den Objekten der visuellen Modalität analog sind. Sie sind gleichzeitig p o t e n t i e l l f o k u s s i e r b a r e E i n h e i t e n , die als Ganzes einer Weiterverarbeitung zugeführt werden, wenn ihnen Aufmerksamkeit zugewandt wird (Wolff, 1977).

Wissensstrukturen und ihre Funktionen bei der Integration und Segmentierung

Integrations- und Segmentierungseffekte lassen sich auch bei der Wahrnehmung hoch strukturierter Ereignisfolgen nachweisen, deren Komplexitätsgrad weit über dem der in Experimenten verwandten Tonfolgen liegt. Ein Beispiel hierfür aus dem Bereich der visuellen Wahrnehmung ist die Wahrnehmung komplexer Handlungsverläufe (Carroll & Bever, 1975).

Die Arbeit von Carroll & Bever (1975) macht deutlich, daß solche Segmentierungen, die Identifikationsprozesse zu erleichtern vermögen, unter realen Umweltbedingungen in hohem Maße vom Wissen des Wahrnehmenden abhängig sind und auf aktive kognitive Prozesse der Organisation zurückzuführen sind. Die Autoren boten kurze Filmsequenzen dar, die aus zwei Teilen bestanden. Der erste Teil war jeweils der Standard und der zweite konnte entweder mit dem ersten identisch sein, eine neue Handlung zeigen oder die Handlung des ersten Teils "paraphrasiert" darstellen. Die Vpn hatten die Aufgabe, die Vergleichssequenz als "alt", "neu" oder "paraphrasiert" zu beurteilen. Die Reaktionszeiten zeigen, daß eine Grenzmarkierung zwischen den beiden Teilen durch einen Filmschnitt die Entscheidungszeit nicht wesentlich beeinflußt, während ein Handlungswechsel die Entscheidungszeit deutlich reduziert, wobei die Kombination beider Segmentierungskriterien sie sogar noch stärker verkürzt, als jede Grenzmarkierung für sich alleine.

Offenbar reicht die durch den Filmschnitt ermöglichte Segmentierung der beiden Filmteile nicht aus, den Unterschied

deutlich zu markieren. Erst wenn eine strukturelle Grenze durch den Wechsel der Handlung eingeführt wird oder sogar mit einer physikalischen Grenzmarkierung zusammenfällt, wird das Erkennen der Ähnlichkeitsrelation zwischen beiden Filmteilen möglich.

Eine physische, zeitliche Strukturierung in Segmente stellt weder eine notwendige noch eine hinreichende Bedingung für eine die Identifikation durch Organisation und Segmentierung unterstützende Wahrnehmung dar, wie wir auch schon beim Phänomen des subjektiven Rhythmus gesehen haben. Notwendige und hinreichende Bedingung ist die Existenz struktureller Grenzen, die der Wahrnehmende nur aufgrund seines Wissens über den wahrzunehmenden Gegenstandsbereich entdecken und für die Identifikation ausnutzen kann. Eine zusätzliche physische Markierung von Segmentgrenzen ist für die Identifikation dann förderlich, wenn sie potentielle strukturelle Grenzen markiert, wie wir bereits bei der rhythmischen Akzentuierung in Tonfolgen gesehen haben (gleiches gilt für die Rolle des Rhythmus in der Sprachwahrnehmung, vgl. Dooling, 1974; Martin, 1972).

Für Integrationen und Segmentierungen unter Verwendung eines spezifischen Wissens des Hörers ist die Sprachwahrnehmung ein eindrucksvolles Beispiel. Bei Sprache handelt es sich um eine hierarchisch strukturierte Folge von Schallereignissen (zur Sprachstruktur vgl. Chomsky, 1957, 1965; Chomsky & Hall, 1968), die vom Hörer eine Segmentierung erfährt, die sich an der Sprachstruktur (teilweise an der Oberflächenstruktur, teilweise aber auch an der Tiefenstruktur) orientiert (einen Überblick geben Fodor, Bever & Garrett (1974) sowie Hörmann (1977)).

Die von Ladefoged & Broadbent (1960) zum Nachweis von Segmentierungsprozessen bei der Sprachwahrnehmung eingeführte experimentelle Methode besteht darin, einer Folge akustischer Signale (meistens Sprachlauten) ein kurzes Klickgeräusch zu überlagern, das vom Hörer in bezug auf die Position der Signale lokalisiert werden soll. In der psycholinguistischen Übertragung dieses experimentellen Paradigmas zum Nachweis der Funktion syntaktischen Wissens bei der Sprachwahrnehmung wurde einem gesprochenen Satz ein Klick überlagert und seine Position in Relation zu den von der Grammatik festgelegten Konstituentengrenzen variiert. Zahlreiche Untersuchungen mit dieser Methode wiesen eine systematische Fehllokalisationstendenz nach, die in zwei Komponenten zu zerlegen ist: 1. wurde eine von der syntaktischen Struktur unabhängige Vorverlagerung des Klicks gefunden und 2. eine systematische Verschiebung der Lokalisation in Richtung auf die Hauptkonstituentengrenzen eines Satzes, die auch dann in gleicher Größenordnung auftritt, wenn keine prosodischen Cues (Pausen und Akzentvariationen) vorhanden waren (Bever, 1973; Bever, Lackner & Kirk, 1969; Bever, Lackner & Stolz, 1969; Fodor & Bever, 1965; Garrett, Bever & Fodor, 1965).

Dieser zweite Befund wird als experimenteller Beleg dafür interpretiert, daß zwischen syntaktischen Grenzen eine Integration kleinerer Einheiten zu syntaktischen Einheiten höherer Ordnung stattfindet, wobei nicht-integrierte Informationen (nämlich die Klicks) subjektiv an der Stelle wahrgenommen werden, an der die Integration auf einer syntaktischen Ebene abgeschlossen ist.

Gregory (1978) hat ein derartiges Klicklokalisationsexperiment mit kurzen Tonfolgen durchgeführt, in denen strukturelle Grenzen leider nur auf einer einzigen Ebene vorkamen. Er konnte im Prinzip die Befunde der psycholinguistischen Experimente dieses Typs replizieren. Eine Interpretation seiner Befunde als ein Beleg für die Anwendung musikalisch-syntaktischen Wissens bei der Integration und Segmentierung von Tonfolgen ist allerdings nicht möglich, da das einzige "syntaktische" Wissen, über das seine Vpn verfügten, in der Kenntnis der physischen Grenzposition zwischen Tongruppen bestand, das ihnen durch visuelle Darbietung des Notentextes vermittelt wurde.

Daß Wissen über musikalische Strukturen für die kognitive Verarbeitung von Musik von großer Bedeutung sein kann, belegen Befunde von Sloboda (1974, 1976, 1977) zur Größe der Blick-Greifspanne beim Klavierspielen vom Blatt. Während des Spielens unbekannter Musik vom Blatt wurde zu für die Vpn unvorhersehbaren Zeitpunkten der Notentext entfernt und die Vpn gebeten, soweit im Spiel fortzufahren, wie ihnen dies möglich war. Dabei zeigte sich, daß die Klavierspieler immer nur bis zu einer musikalischen Phrasengrenze auswendig weiterspielen konnten. Hierin sehen wir einen deutlichen Beleg für die an strukturellen Grenzen orientierte Segmentierung bei der Wahrnehmung des Notentextes und vermuten, daß sich auch beim Hören von Musik derartige Segmentierungseffekte demonstrieren lassen sollten.

Wenn Integration und Segmentierung hierarchisch strukturierter Reizmuster Einheiten konstituiert, die ihrerseits wieder zu Ein-

heiten höherer Ordnung bis zu einer übergeordneten prosodisch und/oder strukturell definierbaren Segmentgrenze integriert werden, dann durchläuft die Enkodierung einer hierarchischen Struktur selbst eine Hierarchie von Repräsentationen von kleineren Einheiten zu größeren Einheiten und bei der Dekodierung von größeren zu kleineren. Angeregt durch Lashley's (1951) Argumente für eine hierarchische Konzeption der Koordination komplexer Handlungsabläufe und stimuliert durch den offensichtlichen Erfolg linguistischer und psycholinguistischer Modelle, die von der Hypothese einer hierarchischen Form der Repräsentation von Sprache ausgingen, entstanden Ende der sechziger Jahre Modelle des seriellen Lernens und der kognitiven Repräsentation serieller Muster, die von assoziationstheoretischen Erklärungsprinzipien Abschied nahmen (wenn man einmal vom Modell von Estes (1972a) absieht).

Pioniere auf diesem Gebiet waren Johnson (1969, 197o, 1972, 1978a,b; Marmurek & Johnson, 1978) und Miller (1956a,b). Während Johnson's Modell (Johnson, 1978a,b) die Frage nach der inhaltlichen Spezifizierung der Kodierungen offen läßt, gehen die meisten anderen Modelle davon aus, daß es sich - wie zuerst auch für die Sprache angenommen wurde (Fodor, Bever, Garrett, 1974); Weiterentwicklungen in Richtung auf semantisch zentrierte Modelle vgl. Engelkamp (1974) und Hörmann (1976) - um die Kodierung von Oberflächenstrukturregeln handeln müsse (Deutsch, 198ob; Egan & Greeno, 1974; Hacker, 1977; Huybrechts, 1977; Jones, 1974, 1978; Martin, 1972; Restle, 197o, 1974; Restle & Brown, 197o; Simon, 1978).

Alle diese Modelle wurden durch Befunde aus Experimenten mit visuell dargebotenem Reizmaterial unterstützt, wenn man einmal von der gerade erst erschienenen Arbeit von Deutsch (198ob) absieht. Obwohl Analogien zur auditiven Mustererkennung naheliegen (Jones, 1978; Restle, 197o), gibt es gegenwärtig nur diese eine empirische Untersuchung zur Frage der Repräsentation von Tonfolgen im Gedächtnis und keinerlei theoretische Ansätze oder gar empirische Untersuchungen zur Frage der Funktion eines musikalisch-syntaktischen Wissens für die Prozesse, die an der Herstellung einer musikalischen Repräsentation beteiligt sind.

2.3 ZUSAMMENFASSUNG UND FOLGERUNGEN

Ausgangspunkt für diesen Überblick über Phänomene und experimentelle Befunde zur auditiven Mustererkennung war die Frage, ob es möglich sei, Aussagen über die Prozesse der kognitiven Verarbeitung beim Hören von Musik zu machen, ohne dabei zu reflektieren, welches eigentlich diejenigen Strukturen akustischer Reizfolgen sind, die ein Spezifikum der Musik sein könnten. Wir charakterisierten die theoretische Grundposition des Informationsverarbeitungsansatzes als einen solchen, der eine Ausklammerung dieser Frage zu erlauben scheint, weil die Verarbeitung komplexer Reizfolgen auf die Verarbeitung elementarer Reize und ihrer Merkmale reduziert wird und damit für Zufallskombinationen von Tönen wie für epochale Schöpfungen eminenter Komponisten gleich ist.

Die Tragfähigkeit dieses Ansatzes erwies sich als unterschiedlich groß. Sowie die Wahrnehmungsbedingungen nur das sukzessive Hören einzelner, merkmalsarmer Reize gestatten, können die meisten Modelle zufriedenstellende Erklärungen geben. Prozeßmodelle, wie die von Neisser & Hirst (1974) und Warren (1974a), können zwar durch Verweis auf unterschiedliche Verarbeitungsprozesse erklären, warum schnelle Tonfolgen schwieriger als langsame zu identifizieren sind, eine eventuelle Beteiligung aller anderen Merkmale (Tonhöhe, Rhythmus, Intervallgröße und Konfigurationen) am Identifikationsprozeß entgeht ihren Modellen jedoch völlig.

Deutsch's (1969, 1975a,b) Modell der Repräsentation von Tönen und Intervallen im Gedächtnis erwies sich in der Folge als zu einfach zur Erklärung des Hörens und Behaltens von Tonfolgen, da die ihm zugrundeliegende Annahme der Konstruktion von Gedächtnisrepräsentanzen aus einzelnen Tönen und ihren sukzessiven Intervallen der experimentellen Überprüfung nicht standhielt (Cuddy & Cohen, 1976; Davies & Jennings, 1977), weil der weitere tonale Kontext, insbesondere die durch die Töne gebildeten Konfigurationen unberücksichtigt blieb.

Untersuchungen zu denjenigen Merkmalen, auf die sich eine Mustererkennung beim Hören von Tonfolgen stützt, erbrachten zwar Befunde, die in sich konsistent waren und uns zur Aufstellung einer Taxonomie elementarer und komplexer Merkmale tonaler Muster nach der Schwierigkeit ihrer Identifizierbarkeit veranlaßten, nutzten aber die Gelegenheit nicht, theoretische Modelle der Mustererkennung auf diesen Befunden aufzubauen (z.B. Divenyi & Hirsh, 1974; Kerivan & Carey, 1976).

Der Überblick über Organisationseffekte beim Hören von Tonfolgen zeigt vielfältige Formen der Interaktion von Merkmalen bei der Bildung von Gruppen innerhalb einer Tonfolge, die als Gedächtniseinheiten im Sinne des Chunking-Konzepts von Miller (1956a) betrachtet werden. Über einen bloßen Verweis auf die Chunking-Konzeption hinausgehende funktionale Erklärungsmodelle konnten wir nicht anführen.

Ein Teil der berichteten Befunde läßt die Grundannahme der linearen Informationsverarbeitungsmodelle, die Wahrnehmung sei durch eine sukzessive Konstruktion einer Repräsentation eines Wahrnehmungsgegenstandes aus elementaren Einheiten zu erklären, ziemlich unwahrscheinlich erscheinen. Es sprechen einige Ergebnisse, insbesondere die zur Identifikation tonaler Strukturen, dafür, daß so etwas wie die Gestalt, die Konfiguration oder der Konturverlauf vor den sie konstituierenden Elementen repräsentiert sein muß, denn nur so läßt sich die Aussage von Divenyi & Hirsh (1974) und Royer & Garner (197o), daß die Wahrnehmung der Tonfolgen ganzheitlich sei, funktional interpretieren.

Endgültig an die Grenzen des linearen Informationsverarbeitungsansatzes stoßen wir beim Problem der funktionalen Bedeutung musikalisch-syntaktischen Wissens bei der Segmentierung und kognitiven Repräsentation hierarchisch strukturierter Melodien. Das Niveau der theoretischen Behandlung dieses Problems geht nicht über Analogien zur Sprachwahrnehmung hinaus. Wenn musikalisches Wissen in die kognitive Analyse beim Hören von Musik eingreift, kann es sich hierbei nur um Prozesse handeln, die von abstrakteren Ebenen der

kognitiven Repräsentation zu reiznäheren Repräsentationsformen verlaufen und mit den ihnen zeitlich in der Verarbeitungsreihenfolge vorgeordneten Prozessen in Interaktion treten. Wenn wir diesen Aspekt thematisieren, werden wir von der Linearitätsannahme Abstand nehmen müssen, und gleichzeitig werden wir gezwungen sein, uns wieder auf die Beantwortung der Frage einzulassen, welches die spezifisch musikalischen Strukturen sind, auf die der Hörer bei der Wahrnehmung von Musik seine Aufmerksamkeit richtet und über die er ein Wissen verfügt, das ihm überhaupt eine Identifikation komplexerer musikalischer Zusammenhänge gestattet.

Der Überblick über die Modelle und Befunde zur auditiven Mustererkennung hat, so meinen wir, deutlich werden lassen, daß sich ein Modell der Informationsverarbeitung für die Prozesse des Erkennens von Strukturen in Musik und ihrer kognitiven Repräsentation in der Literatur bisher nur in Ansätzen abzeichnet.

Auf dem Weg zu einem die vorliegenden Befunde und theoretischen Ansätze integrierenden Modell sind Lösungen für die drei folgenden Probleme zu erarbeiten: 1. Eine funktionale Erklärung der differentiellen Schwierigkeiten bei der Identifikation der Struktur selbst einfacher Tonfolgen ist bislang noch nicht möglich. Dieses Problem wird im Zentrum der theoretischen Überlegungen und Experimente von Teil I stehen. 2. Mit dem ersten Problem verbunden ist die Frage nach den Formen der Repräsentation musikalischer Strukturen.

Sie wird angesichts der Ganzheitlichkeit der Wahrnehmung von Tonfolgen kaum im Rahmen eines linearen Informationsverarbeitungsmodells zu beantworten sein. In Teil I konzentrieren wir uns weiterhin auf das Repräsentationsproblem kurzer Tonfolgen, werden aber in Teil II das in Teil I entwickelte Modell auf echte musikalische Strukturen übertragen und einer experimentellen Überprüfung unterziehen.

3. Nur im unmittelbaren Zusammenhang mit dem Repräsentationsproblem ist die Frage nach der Struktur und Funktion eines musikalisch-syntaktischen Wissens für die Prozesse der Identifikation musikalischer Strukturen zu beantworten. Auch hierzu versuchen wir in Teil II eine theoretische Lösung, die in einigen ihrer Aspekte experimentell überprüft werden soll.

3. STRUKTURELLE UND FUNKTIONALE ASPEKTE DER REPRÄSENTATION KOMPLEXER REIZMUSTER

Wir verlassen nun vorübergehend den Bereich der auditiven Mustererkennung und wenden uns vergleichbaren theoretischen Problemen der visuellen Wahrnehmung zu. Ausgehend von der grundlegenden Überzeugung, daß sich die Verarbeitung von durch unterschiedliche sensorische Kanäle aufgenommenen Informationen im Prinzip nicht unterscheidet (vgl. Gibson, 1966; Marks, 1978; Neisser, 1967, 1976), erhoffen wir uns davon Anhaltspunkte für eine theoretische Lösung der gerade formulierten Probleme der auditiven Mustererkennung. In Kapitel 4 kehren wir dann mit neuem theoretischen Rüstzeug zu den speziellen Problemen der auditiven Mustererkennung zurück.

Neuere Tendenzen in der Forschung zur visuellen Wahrnehmung deuten darauf hin, daß einige Einsichten der Gestaltspsychologie gegenwärtig wieder aktuell werden (Cooper, 1980), und zwar interessanterweise gerade in den USA, wo ihre Beiträge zur Wahrnehmungspsychologie am stärksten Kritik herausforderten (z.B. Gregg, 1932) und seit Ende der 40-iger Jahre beinahe vollständig in Vergessenheit gerieten.

Die Behauptung von Cuddy & Cohen (1976), Divenyi & Hirsh (1974, 1975) und Royer & Garner (1970), daß für die Musteridentifikation in Tonfolgen konfigurative Merkmale oder die Wahrnehmung einer Gestalt von größerer Bedeutung sei, als die Identifikation isolierter Elemente des Reizmusters, mag zwar insofern für den Bereich der auditiven Wahrnehmung neu sein,

als sie durch experimentelle Demonstrationen unterstützt werden konnte, wiederholt allerdings nur die gestalttheoretische Auffassung, daß die Gestaltgesetze der Organisation - wie z.B. die der guten Fortsetzung, der Nähe und der Geschlossenheit - zum Ausdruck bringen, daß es nicht die Eigenschaften der Elemente, sondern die ihnen in ihrer Gesamtheit zukommenden konfigurativen Merkmale sind, die die Wahrnehmung von Objekten und Ereignisfolgen bestimmen (Koffka, 1935; Metzger, 1963).

Diese Sichtweise ist der der Protagonisten des Informationsverarbeitungsansatzes, der sich auf eine Analyse unabhängiger Merkmale (features) stützt (klassische Ansätze: Deutsch & Deutsch, 1963; Hebb, 1949; Selfridge, 1959; Sutherland, 1959; Treisman, 1964; modernere Varianten: z.B. Massaro, 1975; Rumelhart, 197o; Tversky, 1977), insofern diametral entgegengesetzt, als von ihnen behauptet wird, die Wahrnehmung des Ganzen (eines Objekts oder eines Ausschnitts aus der Umwelt) werde durch konstruktive Prozesse aus der Analyse elementarer Merkmale erschlossen. Den Ansatz der linearen Informationsverarbeitungsmodelle kann man also als a n a l y t i s c h im Gegensatz zu dem g a n z h e i t l i c h e n der Gestaltpsychologie charakterisieren.

Argumente für die auch im Zusammenhang des Musikhörens phänomenal unmittelbar einsichtige ganzheitliche Betrachtung der Wahrnehmung werden wir bei einem kurzen Streiflicht auf die Gestalttheorie der Wahrnehmung skizzieren, die der Musik einige überzeugende Beispiele für die Wirkung von Gestaltgesetzen zu verdanken hat. Anschließend werden wir experimentelle Belege und Ansätze für funktionale Modelle der visuellen

Informationsverarbeitung diskutieren, die die Hypothese der Ganzheitlichkeit der Wahrnehmung unterstützen.

Einem Modell der Wahrnehmung und Repräsentation musikalischer Strukturen nähern wir uns in einem ersten Schritt, indem wir neuere Modelle der visuellen Wahrnehmung und Repräsentation komplexer Reizmuster diskutieren. Den Kristallisationspunkt dieser Diskussion bildet das Konzept einer s t r u k t u r e l l e n u n d f u n k t i o n a l e n H i e r a r c h i e v o n R e p r ä s e n t a t i o n s e b e n e n, wie es von Wolff (1977) konzipiert und von Neumann (198o) weiter ausgearbeitet wurde.

3.1 GESTALTTHEORIE UND MUSIKWAHRNEHMUNG

Schon Mach (1886) hat die Unmittelbarkeit des musikalischen Erlebens durch den Begriff der "Tongestalt" charakterisiert und damit zum Ausdruck gebracht, daß offenbar keine "Elementarempfindungen" zur Wahrnehmung einer Melodie zusammengefügt werden müssen, wie es noch der damals gängigen Auffassung der Elementenpsychologie entsprach (Lipps, 19o2; Meyer, 19oo; Seashore, 1938; Stumpf, 1883/189o). Diese "Tongestalt" betrachtete Mach (1886) allerdings noch als Empfindung, wogegen Ehrenfels (189o) einwandte, daß nur Gegenwärtiges, nicht aber etwas von der Dauer einer Melodie als Empfindung bezeichnet werden könne (vgl. Witte, 1952). Darüber hinaus betont Ehrenfels (189o), "die Melodie sei, phänomenologisch betrachtet, nicht eine bloße Zusammenfassung der Töne, sondern etwas Neues, das zwar mit jener Zusammenfassung zugleich,

aber doch unterscheidbar von ihr vorliege" (Witte, 1952, S. 456).

Auch die Gestaltkriterien der "Übersummativität" und "Transponierbarkeit" erläutert Ehrenfels mit musikalischen Beispielen: Ein wahrgenommener Ton besteht zwar physikalisch gesehen aus einer Summe von Partialtönen, aber der resultierende Klangfarbeneindruck ist etwas qualitativ anderes, als die Summe einzelner Tonempfindungen; eine Melodie besteht zwar aus einzelnen Tönen, die Wahrnehmung einer Melodie ist aber zugleich mehr als die Summe der Wahrnehmungen einzelner Töne; und wäre der Eindruck von einer Melodie lediglich die Summe der Eindrücke der betreffenden Elemente, dann müßten die Eindrücke von zwei unterschiedlichen Melodien einander um so ähnlicher sein, je ähnlicher die Elemente der betreffenden Melodien sind, aber trotz vollständiger Verschiedenheit der Elemente transponierter Melodien ist der Höreindruck derselbe (vgl. auch die empirischen Untersuchungen zur Transformation von Melodien, Dowling (1978)).

Gestaltpsychologische Prinzipien haben sich deutlich in den Darstellungen der auditiven Wahrnehmung und Musikpsychologie bei Kurth (1931), Mursell (1937), Revesz (1946) und Vernon (1934) niedergeschlagen. Kurth (1917, 1931), der Schöpfer des Begriffs "Musikpsychologie", sieht die Musik unter energetischen Gesichtspunkten und betont, daß es nicht so sehr der jeweilige Ton sei, der wahrgenommen werde, sondern vielmehr die B e w e g u n g s p h a s e n einer Melodie, die ihre Gestalt ausmachen.

Er hat den Gedanken der Determination der Bedeutung des Einzelnen durch das Ganze nicht nur auf die Melodiewahrnehmung, sondern auch auf die Erklärung von Konsonanz und Dissonanz, Akkordik und Rhythmik angewandt (vgl. de la Motte-Haber, 1976) und damit die Kontextabhängigkeit der musikalischen Höreindrücke, z.B. das "Voraushören", erklären können.

Den Gesichtspunkt der Ganzheitlichkeit der Wahrnehmung von Musik relativieren Crickmore (1972)und Koestler (1964), indem sie darauf verweisen, daß eine Melodie nur als Ganze im Verhältnis zu den ihr untergeordneten "Unterganzen" erscheine, d.h. es hängt von dem Niveau der musikalischen Struktur ab, auf das der Hörer seine Aufmerksamkeit gerade richtet, was als Ganzes wahrgenommen wird: Dies können musikalische Phrasen sein, die als "Unterganze" zu einem Thema integriert werden, es kann sich aber auch um Themen handeln, die im Gesamtzusammenhang einer Exposition eines symphonischen Satzes als Teile in Erscheinung treten.

Demnach ist es also nicht richtig, von d e r Gestalt einer Melodie zu sprechen, sondern sinnvoller, die Wahrnehmung einer musikalischen Gestalt auf das Strukturniveau zu beziehen, auf das die Aufmerksamkeit beim Hören gerichtet ist, wobei diese Orientierung auf ein Strukturniveau während des Hörvorgangs laufend wechseln kann. Musikalische Strukturen sind ähnlich den Kippfiguren: Sie gestatten einen Wechsel von Figur und Grund; aber sie sind gleichzeitig auch vielfältiger als Kippfiguren, da der Grund wieder in sich strukturiert ist und daher eine Ausdifferenzierung zahlreicher Figuren ermöglicht.

Die Beziehung der Herausgliederung musikalischer Gestalten zur Struktur der Musik sehen Crickmore (1972) und Koestler (1964) als hierarchische Verschachtelung. Dieser Ansicht ist auch die der Gestalttheorie verbundene Musikästhetikerin Lissa (1975), die ihre Analyse von Ebenen des musikalischen Verstehens durch die Wahrnehmung struktureller Ebenen eines Musikstücks begründet: Auf der untersten Ebene werden die physischen Charakteristika einzelner Klänge wahrgenommen, auf der zweiten Ebene "größere Ganzheiten ...; es sind ... Relationen der Tonhöhe (Intervalle), Rhythmen, metrische Beziehungen" (Lissa, 1975, S. 67) und auf der dritten Ebene werden syntagmatische Beziehungen wahrgenommen, die größere Abschnitte des musikalischen Verlaufs integrieren. Auf jeder dieser Ebenen (außer der untersten) werden musikalische "Gestalteinheiten" (Lissa, 1975) wahrgenommen, die auf der ihnen jeweils übergeordneten Ebene zu Gestalteinheiten höherer Ordnung entsprechend der zwischen ihnen wahrgenommenen Relationen integriert werden.

Diese Beschreibung erinnert an Meinongs "Logik der Relationen" (Meinong, 1891, 1899), in der die Glieder, zwischen denen eine Relation besteht, als "Fundamente" fungieren (Witte, 1952). Katz (1948) wendet sich gegen eine Gleichsetzung von Gestalt und Relation: "Das Erfassen der zwischen ihren Teilen oder zwischen den Elementen der sie fundierenden Reize bestehenden Beziehungen macht nicht das eigentliche Erfassen der Melodie aus" (Bimberg, 1954, S. 738).

Den Relationsbegriff verwendet Bimberg (1954) ausschließlich zur Kennzeichnung der Beziehung eines einzelnen Tons zum Grund-

ton (Grundtonbezogenheit), zum Leitton oder zur Dreiklangstruktur, während er den Gestaltbegriff im Sinne der Einbettung einer Tonfolge in einen sie vereinenden Bewegungsverlauf verwendet.

Wenn wir Bimberg's engen Relationsbegriff etwas verallgemeinern, dann können wir folgendes sagen: Zwei Melodieteile (z.B. zwei Takte), zwischen denen eine Relation besteht (z.B. die eines Kontrastes: gegenläufige Melodiekontur), können, wenn auf sie die Aufmerksamkeit gerichtet wird, als zwei sukzessive Gestalten (Melodieverlaufsfiguren) gehört werden; wird dagegen die Aufmerksamkeit auf das übergeordnete Strukturniveau gerichtet (Subphrasenniveau z.B. innerhalb einer achttaktigen Periode), dann werden dieselben Takte als eine sich in der Zeit entwickelnde Gestalt gehört, wobei die Relation der Takte nicht als solche repräsentiert ist, sondern primär die sie zusammenschließende Melodieverlaufsfigur; ihre Relation ist allerdings im nachhinein aus der wahrgenommenen Verlaufsfigur rekonstruierbar.

Was wir hier mit Termini der Gestaltpsychologie beschrieben haben, ist bereits in nuce ein Modell der kognitiven Repräsentation musikalischer Strukturen beim Hören von Musik. Wenn wir dieses in operationalisierbare Termini eines funktionalen Informationsverarbeitungsmodells übersetzen könnten, wären wir bereits einen großen Schritt weiter in Richtung auf eine theoretische Beantwortung der Frage nach der Repräsentation musikalischer Strukturen.

Die Gestaltprinzipien der Gruppierung bleiben allerdings häufig nichts weiter, als eindrucksvolle Demonstrationen möglicher Organisationseffekte (Hochberg, 1974). Was bis vor kurzem noch fehlte - und für den Bereich der auditiven Wahrnehmung noch heute fehlt - ist eine experimentelle Überprüfung der gestalttheoretischen Aussagen im Rahmen des Informationsverarbeitungsansatzes, der theoretische Aussagen über die Prozeßabfolge und die Verarbeitungsmechanismen im Kontext unterschiedlicher situativer Anforderungen erlauben würde. Eine Ausnahme bildet hier Garner mit seinen Mitarbeitern (Garner, 1974), der nachweisen konnte, daß die Variable "pattern goodness" kognitive Leistungen (Diskrimination, Reproduktion, Wiedererkennen) in einer Vielzahl visueller und auditiver Aufgaben systematisch beeinflußt. Diese Arbeiten sowie die von Attneave (1954) und Hochberg (1968; Hochberg & Brooks, 196o; Hochberg & McAlister, 1953) untersuchten zwar gestaltpsychologische Organisationsprinzipien mit quantitativen Methoden, testeten aber keine Hypothesen über funktionale Aspekte der Informationsverarbeitung. Ihr Beitrag ist aber insofern von Bedeutung, als er die empirische Basis für die Wirksamkeit der Gestaltgesetze der Wahrnehmung stark erweitert hat.

Welche Konsequenzen lassen sich aus dieser Beschreibung der Musikwahrnehmung im Rahmen eines prozeßdynamisch ausgerichteten Informationsverarbeitungsansatzes ziehen? Aus Meinong's (1891, 1899) noch in der Elementenpsychologie verhafteten Sicht der Fundierung von Gestaltqualitäten durch ihre Elemente ließe sich vielleicht ableiten, daß die Verarbeitung der musikalischen Elemente - wie auch immer man sie definieren mag - zeitlich vor der kognitiven Repräsentation einer Melodie geschehen müßte, was sich in einem

linearen Modell, das auf einer isolierten Analyse von Merkmalen aufbaut und durch konstruktive Prozesse eine Repräsentation einer ganzen Melodie erstellt, darstellen ließe.

Neisser (1967) kommt unter Berücksichtigung gestaltpsychologischer Gesichtspunkte der Figurwahrnehmung (Asch, 1962; Asch, Ceraso & Heimer, 196o) zu der umgekehrten Hypothese: Eine globale Stufe der Analyse müsse der Herauslösung von Einzelheiten vorausgehen, d.h. das Ganze werde zeitlich vor seinen Teilen verarbeitet. Experimentelle Untersuchungen, die Neisser's (1967) Version unterstützen, wurden erst in den letzten Jahren durchgeführt.

3.2 VERARBEITUNG VON STRUKTUREN UND KONFIGURATIONEN

Wenn wir nun von einer überwiegend phänomenologisch ausgerichteten Betrachtungsweise zu einer funktionalen übergehen, können wir nicht das bei der phänomenologischen Sicht im Vordergrund stehende Problem der Repräsentation des Wahrgenommenen isoliert betrachten, sondern müssen vielmehr das Problem der Bildung von Repräsentationen thematisieren, also Wahrnehmungsprozesse und Repräsentationsstrukturen als in vielfältiger Hinsicht voneinander abhängig betrachten: "It is axiomatic within an information-processing framework that one cannot discuss representation without considering processes" (Palmer, 1978a, S. 265).

Für Neisser's (1967) Hypothese der Dominanz einer ganzheitlichen Wahrnehmung über die Wahrnehmung von Elementen gibt es inzwischen einige überzeugende experimentelle Belege. Wir werden aber sehen, daß Neisser's Hypothese noch einer Modifikation bedarf, um den

Prozessen der Wahrnehmung hochstrukturierter Reize in einer selbst wieder strukturierten Umgebung gerecht zu werden.

Effekte der Konfiguration bei der visuellen Identifikation

Wenn die Wahrnehmung bei der Detektion elementarer Merkmale (features) beginnen und hieraus durch konstruktive Prozesse eine Repräsentation eines wahrgenommenen Objekts erstellen würde, dann müßten sich die Bedingungen für die Wahrnehmung des Objekts vollständig auf die Anwesenheit oder Abwesenheit bestimmter, das Objekt typischerweise charakterisierender Merkmale zurückführen lassen. Solange Einzelreize auf einem homogenen Hintergrund dargeboten werden, lassen sich die Prozesse mit diesem Modell noch einigermaßen zufriedenstellend beschreiben. Ist der Reiz aber in sich strukturiert und/oder wird er auf einem strukturierten Hintergrund dargeboten, der selbst wieder als strukturierte Menge von Reizen betrachtet werden kann, dann zeigen sich kontextabhängige Identifikationsleistungen, die nur durch ein Modell erklärt werden können, das davon ausgeht, daß größere Einheiten als individuelle Merkmale für die Verarbeitungsmechanismen bei der Wahrnehmung eine wichtige Rolle spielen.

Vergleicht man die Identifikation von Wörtern und Buchstabenfolgen unterschiedlichen Redundanzgrades, dann erhält man eine höhere Identifikationsrate und -geschwindigkeit für Wörter als für Buchstabenfolgen und bei Buchstabenfolgen einen Effekt des Redundanzgrades, der in einer besseren Identifizierbarkeit bei höherem Redundanzgrad besteht (z.B. Mewhort, 1966, 1972; Miller, Bruner & Postman, 1954; Wolff, 1977). Diese und andere Befunde zum Lesen

haben zu der Hypothese geführt, daß größere Einheiten als einzelne Merkmale von Buchstaben oder isolierte Buchstaben bei der visuellen Identifikation von Wörtern verarbeitet werden (Broadbent, 1977; Gibson, 1969; Spoehr & Smith, 1979; Wolff, 1977).

Die Arbeiten von Biederman (1972; Biederman, Glass & Stacey, 1973; Biederman, Rabinowitz, Glass & Stacey, 1974) und Reinert (1980) über die Wahrnehmung von Objekten in realen Umweltszenen zeigen, daß die Identifikation in starkem Maße von der Organisation, in die das Objekt eingebettet ist, abhängt.

Selbst dann, wenn nur ein einziges Feature kritisch für eine Identifikation ist, zeigen sich derartige Organisationseffekte. In Weisstein's Versuchen zum "Objekt-Überlegenheitseffekt" (Weisstein & Harris, 1974; Williams & Weisstein, 1978) erwies sich die Identifikationsrate für Linienorientierungen in einem visuellen Kontext als höher, wenn dieser Kontext ein kohärentes Objekt bildete und dreidimensional erschien, als wenn der Kontext kein Objekt bildete und nur zweidimensional abgebildet war. Ein Strich ohne Kontext war sogar schwieriger als einer mit Kontext zu identifizieren.

Eine auditive Analogie zum "Objekt-Überlegenheitseffekt" stellt der Kontexteffekt von Dewar, Cuddy & Mewhort (1977) dar. Die Autoren boten den Vpn eine Folge von sieben Tönen dar, der in einer experimentellen Bedingung nach einer kurzen Pause zwei Töne folgen und in einer anderen Bedingung zwei Folgen von sieben Tönen. Aufgabe der Vpn war es anzugeben, welcher der beiden Vergleichstöne in der Standardtonfolge vorkam bzw. welche der beiden

Vergleichstonfolgen der Standardtonfolge entsprach, wobei genau ein Ton in einer der beiden Vergleichstonfolgen von der Standardtonfolge abwich. Sowohl für an der Tonalität orientierte Tonfolgen als auch für Zufallsfolgen von Tönen lag die Wiedererkennungsrate für Vergleichstonfolgen höher als für isolierte Vergleichstöne, aber unabhängig davon war die Wiedererkennungsrate für an der Tonalität orientierte Tonfolgen deutlich höher als für Zufallsfolgen. Auch hier zeigt sich also eine Überlegenheit des Erkennens im Kontext, die für organisierte Kontexte noch höher ausfällt als für unorganisierte.

Auch Pomerantz und Mitarbeiter (Pomerantz, 1978; Pomerantz & Garner, 1973; Pomerantz, Sager & Stoever, 1977; Pomerantz & Sager, 1976; Pomerantz & Schwaitzberg, 1975) konnten in einer Serie von Experimenten nachweisen, daß das Erkennen eines Liniensegments von der Konfiguration der das Segment umgebenden Linien abhängt. Die von ihnen verwandten Displays bestanden aus einem Quadrat, in dessen Quadranten drei gleichorientierte Linien und eine anders orientierte Linie enthalten waren. In den experimentellen Bedingungen variierte der die kritische Linie umgebende Linienkontext. Die Aufgabe der Vpn bestand darin, so schnell wie möglich denjenigen Quadranten anzugeben, in dem sich die Orientierung des kritischen Liniensegments von der in den anderen Quadranten unterschied. Die Identifikationszeit für das kritische Liniensegment war am kürzesten in dem Kontext, in dem es nicht in eine geschlossene Figur eingebettet wurde; die nächst höhere Reaktionszeit wurde für kontextfreie Darbietung erzielt und die längste für Kontexte, in denen das Liniensegment Teil einer geschlossenen Figur war.

Ganz ähnliche Experimente führten Banks & Prinzmetal (1976) und Prinzmetal & Banks (1977) durch. Ihre Vpn sollten einen Buchstaben in einem buchstabenähnlichen Kontext identifizieren, der entweder so aufgebaut war, daß er nach den Gestaltgesetzen mit den Kontextelementen gruppiert werden mußte oder aber als isoliertes Element in Erscheinung trat. Reaktionszeiten und Fehlerzahlen waren dann am höchsten, wenn der Buchstabe zum Kontext gruppiert wurde. Die Autoren interpretieren ihre Befunde als unvereinbar mit der Hypothese von Estes (1972), wonach die Erschwerung der Identifikation eines Signals auf einem Hintergrund nicht zu beachtender Reize durch eine Interaktion lokaler Featuredetektoren zu erklären ist. Stattdessen nehmen sie an, daß der Identifikation eine Stufe der Gruppierung vorausgeht, deren Operationsprinzipien mit den Gestaltgesetzen übereinstimmen. In dieser Sicht konfigurativer Effekte stimmen sie mit Bregman (1978a,b) überein, dessen Analyse der Streamorganisationen die Wirksamkeit von Gestaltgesetzen für die auditive Modalität überzeugend belegt hat.

Weitere experimentelle Paradigmen, mit denen die Wirksamkeit konfigurativer Merkmale für die visuelle Wahrnehmung demonstriert wurde, sind die rückwirkende Maskierung (Bell & Handel, 1976; McClelland, 1978), womit gezeigt wurde, daß objekthafte Reize weniger effektive Maskierreize als Zufallsreize und Reize mit schlechter Gestalt sind, die mentale Synthese einzelner Segmente zu einem Gesamtmuster (Palmer, 1977; Thompson & Klatzky, 1978), die für Segmente, die in der Synthese eine gute Gestalt besitzen, schneller durchgeführt werden kann als für Segmente, die zu einer schlechten Gestalt führen, und Gedächtnisexperimente (Bower & Glass, 1976), die zeigen, daß die Reproduktion einfacher

Strichzeichnungen durch "gute" Cues (Teile einer durch Nähe oder gute Fortsetzung organisierten Gestalt) stärker als durch schlechte erleichtert wurde.

Alle diese Befunde sprechen gegen ein Modell der Kodierung visueller Muster durch Featureanalyse (weitere Arbeiten zitiert Broadbent (1977)). Stattdessen müssen Informationsverarbeitungsmodelle, die diesen und den über einem halben Jahrhundert vorher gewonnenen Befunden der Gestaltpsychologie Rechnung tragen wollen, davon ausgehen, daß bei der Wahrnehmung strukturierter Reizmuster Verarbeitungseinheiten funktional wirksam werden, die im Vergleich zu physischen Reizelementen und Einzelmerkmalen Einheiten höherer Ordnung darstellen, deren Wahrnehmung sowohl von ihrer Binnenstruktur als auch von der Struktur des sie umgebenden figuralen bzw. tonalen Kontextes determiniert wird.

Globale und lokale Reizmusteranalyse

Banks & Prinzmetal (1976; Prinzmetal & Banks, 1977) bieten eine funktionale Erklärung für die Konfigurationseffekte bei der Identifikation visueller Muster an, die in der Annahme eines zweiphasigen Analyseprozesses besteht. In der ersten Phase wird ein Strukturierungsmechanismus ("parser") aktiv, der den visuellen Input simultan nach Gestaltprinzipien in Einheiten oder Gruppen zusammengehöriger Elemente aufgliedert. Erst im Anschluß an diese Gliederung erfolgt eine analytische Form der Verarbeitung, die in einer sequentiellen Verarbeitung der ausgegliederten Einheiten oder Gruppen besteht. Die zeitliche Prozeßabfolge läßt sich nach diesem Modell als Übergang von

einer globalen Form der Analyse zu einer lokalen oder als Übergang von einer grobkörnigen, detailarmen Auflösung des Wahrgenommenen zu einer feinkörnigen, detailreichen Auflösung beschreiben ("top-down"-Verarbeitung, vgl. auch Kahneman, 1973).

Diese Hypothese über die Prozeßreihenfolge stimmt mit den Befunden zur Aktualgenese der Wahrnehmung überein (einen Überblick geben Flavell & Draguns, 1957), die zeigen, daß mit zunehmender Expositionszeit mehr Details wahrgenommen werden: Zuerst wird nur der Ort des Objekts lokalisiert, dann folgt eine Differenzierung von Figur und Grund mit einer eher globalen Wahrnehmung der Form und schließlich eine Wahrnehmung der Details der Figur.

Navon's (1977) Experimente stellen eine eindrucksvolle Bestätigung dieser Hypothese dar. In seinem ersten Experiment hörten die Vpn den gesprochenen Buchstaben H oder S und sahen gleichzeitig in 5o % der Fälle ein H oder S auf einem Monitor. Die visuell dargebotenen Buchstaben (globales Niveau) waren entweder aus kleinen Buchstaben desselben oder des jeweils anderen Typs zusammengesetzt (lokales Niveau). Aufgabe der Vpn war es, für jeden gehörten Buchstaben die entsprechende Taste zu drücken. Die Reaktionszeiten zeigen eine deutliche Interferenzwirkung bei Nichtübereinstimmung des gehörten mit dem visuell dargebotenen Buchstaben auf globalem Niveau, aber keine Interferenz aufgrund des lokalen Niveaus. Dieser Befund zeigt, daß lokale Information, die nicht für die Reaktionsentscheidung wichtig ist, auch nicht verwertet wird, d.h. die visuelle Analyse kann nach der globalen Analyse abgebrochen werden.

Wenn es sich hierbei um eine willkürlich steuerbare Verarbeitung handeln würde, sollte die globale Analysestufe übersprungen werden können. Daß dies aber nicht der Fall ist, zeigt Navon's Experiment 3: Während die Buchstaben auf lokalem Niveau für die Identifikation auf globalem Niveau ohne Bedeutung sind, verzögert die Nichtübereinstimmung von globalem und lokalem Niveau die Identifikation der Buchstaben auf lokalem Niveau. Diese Asymmetrie deutet auf eine zeitlich der lokalen Analyse vorgeordneten globalen Analyse hin, die nicht willkürlich übersprungen werden kann.

Ein weiterer Beleg für die Notwendigkeit der Differenzierung zwischen einer globalen oder holistischen Form der visuellen Analyse und einer lokalen oder analytischen Form kommt von Palmer (1978b). Seine Vpn betrachteten Strichzeichnungen von Szenen mit jeweils sechs Objekten, die sie in einer anschließend gezeigten Bilderserie wiedererkennen sollten. Die im Wiedererkennungsdurchgang gezeigten Objekte waren entweder alle identisch oder enthielten ein systematisch transformiertes Objekt. Die Transformationen bestanden in der Veränderung der Größe, im Ersetzen eines Objekts durch eines anderer Identität oder durch Veränderung physischer Merkmale eines Objekts unter Beibehaltung seiner Identität. Same-different-Reaktionszeiten und Augenbewegungen wurden registriert.

Von besonderem Interesse ist hier der Befund, daß die Reaktionszeiten bei Ersetzen eines Objekts durch ein anderes kürzer waren als bei Änderung physischer Merkmale ohne Änderung seiner Identität; hierbei fixierten die Vpn sogar das veränderte oder ersetzte Objekt nicht einmal. In anderen different-Bedingungen wichen die Vpn von ihrer systematischen Suchstrategie ab, die sie bei same-Bedingungen anwandten, und fixierten das veränderte Objekt unmittelbar.

Diese Befunde lassen vermuten, daß eine globale Analyse eines Reizmusters bei grobkörnigen, globalen Veränderungen bereits zu einer reaktionsadäquaten Repräsentation führen kann und daß erst bei relativ feinkörnigen lokalen Veränderungen eine Fixierung der Lokalisation der Veränderung und eine detailliertere Analyse stattfindet, die durch die Ergebnisse der globalen Analyse gesteuert wird (vgl. auch Palmer, 1977).

Broadbent (1977) und Navon (1977) bringen diese Befunde im Zusammenhang mit der Ortsfrequenzhypothese über den Funktionsmechanismus visueller Rezeptoren (siehe hierzu Breitmeyer, 1975; Broadbent, 1975; Campbell, 1974; Ginsburg, 1976; Lupp, Hauske & Wolf, 1976). Nach Henning, Hertz & Broadbent (1975) ist das Hochfrequenzsystem (Detailanalyse) dem Niederfrequenzsystem (Globalanalyse) hierarchisch untergeordnet, d.h. das Niederfrequenzsystem dient der Entdeckung großer Gebiete hochfrequenter Stimulation, die von niederfrequenter Stimulation umgeben sind, aber das hochfrequente System kann das Umgekehrte nicht leisten. Auf diese Weise kann das niederfrequente System simultan mehrere Segmente ausgliedern, auf die nacheinander Aufmerksamkeit gerichtet werden kann. Wenn dies der Fall ist, kann aufgrund des hochfrequenten Systems eine Detailanalyse einzelner Segmente und damit eine Identifikation stattfinden.

Auf ein Beispiel aus dem auditiven Bereich übertragen kann die Aufspaltung einer Tonfolge in zwei Streams auf ein niederfrequentes Dekompositionssystem zurückgeführt werden, dessen Informationen die Fokussierung auf einen Stream gestatten, während das hochfrequente System dem Erkennen von Tonhöhenbeziehungen im fokussierten Stream dient.

Nun kann man allerdings das globale, niederfrequente System und das lokale, hochfrequente System nicht als zwei zeitlich ausschließlich sukzessiv arbeitende Verarbeitungsstufen betrachten, wie dies für die Systemkomponenten der meisten Informationsverarbeitungsmodelle gilt, denn der Output des globalen Niederfrequenzsystems ist nicht zugleich Input für das lokale Hochfrequenzsystem. Vielmehr geben die Befunde von Navon (1977) zur Interferenz zwischen lokaler und globaler Information bei lokaler Analyse erst dann einen Sinn, wenn man annimmt, daß das globale System ständig aktiv ist, um potentielle lokale Analysen zu ermöglichen. Die in den Daten Navon's erscheinende Sukzession stellt nicht notwendigerweise eine zeitliche Sukzession des Analysebeginns der Systeme dar, sicherlich aber eine zeitliche Sukzession der aus der Analyse resultierenden Repräsentationen. Das Ganze wird also nicht unbedingt vor seinen Teilen analysiert, sondern parallel zu ihnen, aber es wird vor seinen Teilen repräsentiert.

Bemerkenswert an Navon's (1977) Modellansatz ist die Annahme zweier Systeme - die Broadbent (1977) als Systeme der präattentiven Verarbeitung bezeichnet - die parallel arbeiten, aber trotzdem eine Abhängigkeitsbeziehung zeigen. Solche Abhängigkeiten sind mit der Linearitätsannahme des Informationsverarbeitungsansatzes nicht vereinbar. Da zwischen der globalen und der lokalen Analyse eine Dominanzbeziehung besteht, die sich in der einseitigen Beeinflussung der lokalen durch die globale Analyse ausdrückt, können sie nur als Teilsysteme einer f u n k t i o n a l e n V e r a r b e i t u n g s h i e r a r c h i e betrachtet werden, die die hierarchischen Beziehungen zwischen Reizaspekten (Einbettung der lokalen Merkmale in die globalen bei Navon (1977), propositionale Hierarchie von figuralen Einheiten bei Palmer (1975, 1977)) auf eine Hierarchie von Repräsentationen abbilden.

Die Idee einer hierarchischen kognitiven Repräsentation der wahrgenommenen Welt ist keineswegs neu; aber erst in den letzten Jahren beginnt sich die Erkenntnis durchzusetzen, daß eine hierarchische Repräsentation nicht nur R e s u l t a t einer Verarbeitung sein kann, sondern gleichzeitig eine d i e V e r a r b e i t u n g s t e u e r n d e F u n k t i o n besitzen muß. Damit wird die von Palmer (1978a) beklagte Loslösung des strukturellen vom funktionalen Aspekt der kognitiven Repräsentation aufgehoben.

3.3 STRUKTURELLE UND FUNKTIONALE ASPEKTE DER KOGNITIVEN REPRÄSENTATION

Das Hierarchiekonzept

In der Psychologie wird das Konzept der Hierarchie als funktionales oder strukturelles Konstrukt verwendet, wenn Dominanzbeziehungen zwischen Verhaltenstendenzen zu erklären sind oder Über- bzw. Unterordnungsbeziehungen zwischen Elementen der Reizmuster zu berücksichtigen oder zwischen Einheiten kognitiver Repräsentationen zu vermuten sind. Im funktionalen Sinne ist die "habit-family hierarchy" bei Hull (1934), die Assoziationshierarchie (z.B. bei Deese, 1965) und die Bedürfnishierarchie Maslow's (1954) zu verstehen; ein strukturelles Konstrukt stellt z.B. die Hierarchie der Wesenszüge bei Guilford (1959) dar sowie die kognitive hierarchische Organisation im Modell der semantischen Netze bei Quillion (1968), in Gedächtnismodellen z.B. bei Johnson (1978a,b) und Thorndike (1977) und Problemlösungsmodellen (vgl. Egan & Greeno, 1974).

Das Konzept der Hierarchie gestattet es, Verhalten gleichzeitig auf mehreren Ebenen zu beschreiben. Was auf einer Ebene den Status einer größeren Einheit besitzt, besitzt auf der darüberliegenden Ebene den Status eines Elements, das integraler Bestandteil einer übergeordnete Zusammenhänge repräsentierenden Einheit ist. Mit dieser Bedeutung des Hierarchiebegriffs ist der Hull'sche Begriff der Gewohnheitshierarchie und der Begriff der Assoziationshierarchie nicht vereinbar. Bei diesen handelt es sich um p r o b a b i l i s t i s c h e Ordnungen nach der Stärke a u s t a u s c h b a r e r Reaktionen in einer Liste, während das Hierarchiekonzept im Sinne einer Ebenengliederung mit unterschiedlicher Korngröße der Beschreibungseinheiten von d e t e r m i n i s t i s c h e n Beziehungen zwischen den Einheiten auf benachbarten Ebenen ausgeht: Die jeweils untergeordneten Einheiten konstituieren die übergeordneten und sind daher nicht a u s t a u s c h b a r , ohne daß gleichzeitig damit die gesamte Organisation einer Hierarchie tangiert würde ("coordinate structure constraint", McCawley, 1974).

Kognitive Hierarchiemodelle wurden bis vor kurzem ausschließlich für die a b s t r a k t e Gedächtnisrepräsentation von Organisationen formuliert. Mit abstrakten Repräsentationen meinen wir hier solche, die als E r g e b n i s konstruktiver Verarbeitungsprozesse betrachtet werden, wie z.B. die Speicherung einer Ereignissequenz bei der Identifikation serieller Muster (vgl. Jones, 1974, 1979; Restle, 197o). Solche ausschließlich strukturell begründeten Modelle können nicht die Ganzheitlichkeit der Wahrnehmung erklären, weil die ganzheitliche Repräsentation, z.B. die Musterrepräsentation auf der höchsten Hierarchieebene, als Ergebnis der integrativen Leistung des Gedächtnisses

verstanden wird, d.h. auch in diesen Modellen ist die Wahrnehmung als Repräsentation elementarer Einheiten konzipiert, aus denen erst in einem "bottom-up"-Prozeß eine Repräsentation des gesamten Wahrnehmungsobjekts konstruiert werden muß (Restle & Brown, 197o). Da sich der phänomenale Eindruck von einem Objekt oder einer Melodie nicht erst nach Beendigung, sondern bereits während des Wahrnehmungsvorgangs einstellt, muß die Herstellung einer ganzheitlichen Repräsentation von Gedächtnissystemen weitgehend unabhängig sein.

Das hierarchische Repräsentationsebenenmodell von Wolff (1977)

Das Problem der Erklärung von Effekten der Reizmusterorganisation für die Prozesse der Wahrnehmung, das sich uns bei der auditiven Mustererkennung stellt, bildet einen Schwerpunkt der Arbeit von Wolff (1977). Wolff stellte sich das Problem bei der Erklärung des Effekts der Redundanzausnutzung tachistoskopisch dargebotener Buchstabenzeilen (z.B. Baddeley, 1964; Mewhort, 1966, 1972; Miller, Bruner & Postman, 1954; Wolff, 1971). Um die Hypothese der Abtastbeschleunigung (Mewhort, 1974) mit der Chunkinghypothese (Baddeley, 1964) in befriedigender Weise integrieren zu können, nahm Wolff an, daß bereits in einer primären Reizrepräsentanz organisierte Chunks vorliegen - also nicht erst auf dem Niveau des Kurzzeitgedächtnisses (Chunks durch Rekodierung) gebildet werden, wie dies der klassischen Chunkinghypothese Miller's (1956a,b) entsprechen würde (s.a. Baddeley, 1964; Mewhort, 1972, 1974), die visuelle Verarbeitungseinheiten darstellen, die als Ganzes dem VIS[1)] entnommen werden können:

---

1) "visual information store", vgl. Sperling (196o); Neisser's (1967) "ikonischer Speicher"

"Wir müssen daher annehmen, daß im redundanten Kontext die reproduktionsrelevante Information sowohl im Chunking-Verband als auch in isolierten Elementrepräsentanzen zur Verfügung steht, während sie bei zufälligen Reizen nur in den Elementrepräsentanzen vertreten ist" (Wolff, 1977, S. 415).

Das visuelle Reizmuster wird also im organisierten Fall als doppelt repräsentiert betrachtet. Um eine Interferenz bei der Informationsentnahme zwischen beiden Repräsentationen theoretisch auszuschliessen, mußte gefordert werden,

"... daß die Position, mit welcher eine Buchstabenkombination als Chunk repräsentiert ist, von den Positionen, mit welchen die diesen Chunk konstituierenden Elemente als Elementrepräsentanzen vertreten sind, verschieden ist ... Die Position einer Verarbeitungseinheit ist von der Korngröße dieser Einheit selbst abhängig. Dies läßt sich verstehen, wenn man die Reizrepräsentanz als eine H i e r a r c h i e von Repräsentanzen auffaßt, welche horizontal wie vertikal gegliedert ist:

Mit v e r t i k a l e r Gliederung ist gemeint, daß die Reizrepräsentanz verschiedene R e p r ä s e n t a t i o n s e b e n e n (-niveaus) erkennen läßt. So besteht die Reizrepräsentanz ... aus mindestens zwei Repräsentationsebenen: Auf der obersten Ebene ist das Reizmuster als Ganzes repräsentiert (Reizmusterrepräsentanz), auf einem unteren Repräsentationsniveau sind ... Einzelrepräsentanzen vertreten ... Die h o r i z o n t a l e Gliederung fällt für verschiedene Repräsentationsniveaus unterschiedlich aus: Mit der Höhe der Repräsentationsebene nimmt die Korngröße der Verarbeitungseinheiten zu, ihre Anzahl nimmt ab.

Mit dem Begriff der Repräsentationsebene läßt sich erklären, wieso die Position einer Verarbeitungseinheit von ihrer Korngröße abhängt: Verarbeitungseinheiten verschiedener Korngröße befinden sich auf verschiedenen Repräsentationsebenen, ihre Positionen unterscheiden sich also auf der vertikalen Dimension der Reizrepräsentanz.

Die Repräsentationshierarchie redundanter Reize unterscheidet sich nun von derjenigen zufälliger Sequenzen darin, daß zwischen dem Niveau der Elementrepräsentanzen und dem Niveau der Reizmusterrepräsentanz die Ebene der Chunkrepräsentanz existiert ..." (Wolff, 1977, S. 415-416).

Wolff's strukturelle Repräsentationshierarchie unterscheidet sich wesentlich von der bisheriger hierarchischer Repräsentationsmodelle: Der Ort der Repräsentationshierarchie ist innerhalb der Verarbeitungssequenz auf das Niveau der Wahrnehmungsprozesse vorverlagert worden, und damit sind unterschiedliche Korngrößen von Repräsentationen bereits w ä h r e n d der Verarbeitung des Reizmusters möglich und nicht erst im Anschluß daran, wie dies in hierarchischen Repräsentationsmodellen impliziert ist, die Organisationseffekte ausschließlich auf längerfristige Gedächtnissysteme (Kurzzeit-, Arbeits- oder Langzeitgedächtnis) zurückzuführen. Dieses "reiznähere" Repräsentationsmodell schließt die Möglichkeit der Repräsentation des Ganzen wie die seiner Teile ein.

Wie aber kann es die Dominanz und den zeitlichen Vorrang einer globalen Repräsentation gegenüber einer lokalen im Sinne Navon's (1977) erklären?

"Die Fokussierung einer Verarbeitungseinheit setzt stets eine visuelle Analyse auf dem jeweils höheren Repräsentationsniveau voraus, da die Position der zu fokussierenden Verarbeitungseinheit nur auf der höher geordneten Ebene durch simultane Analyse ausfindig gemacht werden kann. Das Aufsuchen selbst (Zugriff) erfolgt aber auf dem niederen Niveau" (Wolff, 1977, S. 424–425).

Diese Annahme ist logisch begründet: Das Erkennen eines Detailmerkmals setzt ein tragendes Objekt voraus; das Objekt selbst muß also notwendigerweise in irgendeiner Form während des Aufsuchens des Details kognitiv repräsentiert sein - und es ist ja auch phänomenal als solches repräsentiert - und bereits analysiert worden sein, zumindest derart, daß eine Ausgliederung von Details erreicht wird. Genau dies ist auch die Argumentation, die Neisser (1967)

zur Einführung des Begriffs der "präattentiven" Prozesse führte. Nur die Annahme, daß die präattentiven Prozesse ausschließlich auf dem Niveau von Merkmalen operieren, ist angesichts der Befunde von Navon (1977) und derjenigen zur Wahrnehmung konfigurativer visueller Reizmuster nicht mehr haltbar (vgl. auch die Kritik Neumanns (198o) am Konzept der "präattentiven" Prozesse). Vielmehr ist jede nicht gerade fokussierte Repräsentationsebene in dem Sinne "präattentiv", als sie nicht gerade intentional beachtet wird (Neumann, 198o). Wenn ein Hörer den Melodienverlauf fokussiert und ihm dabei eine unsaubere Tonproduktion eines Instruments entgeht, dann kann man sagen, daß die Repräsentationsebene einzelner Tonqualitäten in dieser Situation "präattentiv" ist, weil sich die Aufmerksamkeit auf eine weitaus höhere Ebene mit größerer Korngröße konzentriert.

Im Rahmen von Wolff's Modell läßt sich der Vorgang einer ganzheitlichen Wahrnehmung vor der von Details folgendermaßen begründen:

"Die Annahme, daß die räumliche Fokussierung auf einem mehrere Verarbeitungseinheiten (Repräsentanzen) enthaltenden Niveau eine visuelle Analyse auf der nächst höheren Ebene voraussetzt, führt zu der Konsequenz, daß die kontrollierte Informationsentnahme ein hierarchisch gequantelter Prozeß ist. Eine Fokussierung auf einem unteren Repräsentationsniveau setzt die Durchwanderung aller Repräsentationsebenen höherer Ordnung voraus. Die Position einer Verarbeitungseinheit auf einem unteren Repräsentationsniveau einer in mehr als zwei Ebenen vertikal gegliederten Reizrepräsentanz wird also durch hierarchische Einschachtelung gefunden, wobei auf jeder höher geordneten Repräsentationsebene ein visueller Analyseschritt erforderlich ist. Die Anzahl der zu durchwandernden Ebenen, d.h. die Anzahl der Analyseschritte hängt von zwei Sachverhalten ab:
(a) Die Aufgabenstellung bestimmt das Niveau, auf welchem die reproduktionsrelevante Information in einer zur Bewältigung der Wahrnehmungsaufgabe geeigneten Weise entnommen werden kann.

(b) Die Komplexität der Reizrepräsentanz bestimmt die Anzahl der Repräsentationsebenen, die zu durchwandern bzw. zu analysieren sind.

Hieraus folgt, daß die Information jeweils auf der höchsten die reproduktionsrelevante Information in geeigneter Weise enthaltenden Repräsentationsebene entnommen wird" (Wolff, 1977, S. 425-426).

Die funktionale Begründung einer derartigen Repräsentationshierarchie ist aber erst dann vollständig, wenn der Begriff der Verarbeitungseinheit präzisiert werden kann und hierdurch gleichzeitig begründet werden kann, warum eine Verarbeitungseinheit nur als Ganze und nicht gleich die sie konstituierenden Elementinformationen entnommen werden können. Wolff (1977) begründet dies durch Verweis auf die gemeinsame Verbundenheit der Elementinformationen mit einem Objekt an einem Ort:

"Wir nehmen daher an, daß die die Eigenschaften eines Objekts repräsentierenden Merkmale in solcher Weise integrativ verbunden sind, daß sie nicht isoliert, sondern nur als Merkmalsverband entnommen werden können ... Die Verarbeitungseinheit (Repräsentanz) ist also als ein Merkmalsverband, in welchem jede Merkmalsdimension mit einer Ausprägung vertreten ist, hinreichend definiert.

Wir glauben nun, daß die Gebundenheit der Merkmale dadurch gewährleistet ist, daß sie ein und demselben Ort im visuellen Repräsentationsraum ... zugeordnet sind ... Durch Fokussierung dieser Position werden nun die Merkmale als integrale Bestandteile ihres Verbandes entnommen. Sie werden als Eigenschaften eines b e s t i m m t e n Objekts identifiziert, weil sie als Merkmal e i n u n d d e s s e l b e n Verbandes repräsentiert sind" (Wolff, 1977, S. 428-429).

Im Unterschied zu linearen Verarbeitungsmodellen, in denen die Repräsentation des Wahrgenommenen als durch konjunktive Verknüpfung einzelner Features konstruiert gedacht ist, ändert sich

der Inhalt einer Repräsentation im Wolff'schen Modell nicht, sondern es ändert sich die Art der Repräsentation (Neumann, 198o). Durch konjunktive Featureverknüpfung würde eine Wahrnehmung erst allmählich vervollständigt (diese Annahme liegt auch weiterhin dem "levels of processing"-Ansatz von Craik & Lockhart (1972) zugrunde), so daß sich "frühe" und "späte" Repräsentationen inhaltlich unterscheiden müßten. Dem widerspricht aber die empirische Tatsache, daß die sogenannte "späte" semantische Repräsentation eines Wortes tatsächlich aber zeitlich v o r den individuellen Einheiten (Phoneme, Silben, Buchstaben) repräsentiert ist (vgl. Gardiner, 1974; Marmurek & Johnson, 1973). Im Wolff'schen Modell sind dagegen alle Merkmale auf jeder Repräsentationsebene in einer Verschachtelung von Merkmalsverbänden enthalten, aber je nach Höhe des aktiven Repräsentationsniveaus direkt oder nur vermittelt über darunterliegende Niveaus verfügbar.

Dies bringt uns zu einem Punkt, den Wolff (1977) nicht thematisiert hat, der aber für die Frage der ganzheitlichen Repräsentation, speziell der Repräsentation konfigurativer Merkmale von Bedeutung ist. Wenn auf jeder Repräsentationsebene der Inhalt der Repräsentation derselbe und nur die Art der Repräsentation eine andere ist, dann sollte eine erschöpfende Beschreibung eines komplexen Reizes durch Aufzählung seiner Merkmale der niedrigsten Repräsentationsebene möglich sein, oder pointiert formuliert, das Ganze wäre als hierarchische Gruppierung seiner nach ihrer Korngröße geordneten Merkmale aufzufassen. Das Problem besteht nun darin, daß nach dieser Konzeption ein Merkmalsverband nur durch den gemeinsamen Ort der sie konstituierenden Elemente definiert wäre und nicht noch zusätzlich durch ihre r ä u m l i c h e K o n f i g u r a t i o n

zueinander. Die referierten Befunde zur Wahrnehmung konfigurativer Merkmale visueller Muster haben nachdrücklich die Wirksamkeit gestaltpsychologischer Organisationsprinzipien bestätigt. Der Begriff der funktionalen Verarbeitungseinheit muß über seine Bestimmung bei Wolff (1977) hinaus also noch dadurch präzisiert werden, daß die in einem Merkmalsverband repräsentierten Merkmale nicht nur am selben Ort des Wahrnehmungsraums lokalisiert werden, sondern darüber hinaus auch noch qualitativ neue Konfigurationen aufweisen, die auf darunter liegenden Ebenen nicht repräsentiert sein können. Oder anders formuliert: Die in einem Merkmalsverband repräsentierten Merkmale konstituieren nur deshalb ein Objekt und werden dadurch am selben Ort des Wahrnehmungsraumes lokalisiert, weil ihre Konfigurationen den Kriterien einer guten Gestalt genügen.

Wir wollen es zunächst bei dieser recht globalen Formulierung belassen, werden aber später bei der Übertragung dieses Modells auf die auditive Mustererkennung zu einer weiteren Präzisierung gelangen. Es bleibt aber festzuhalten, daß es offenbar nicht genügt, davon auszugehen, daß auf allen Repräsentationsebenen dieselbe Information in nur quantitativ anderer Art repräsentiert ist, vielmehr muß angenommen werden, daß mit zunehmender Höhe des Repräsentationsniveaus zunehmend komplexere und ihrer Korngröße nach räumlich bzw. im auditiven Fall zeitlich weitergespannte konfigurative Merkmale repräsentiert sein müssen. Diese Auffassung ist mit Palmer's (1977) hierarchischem Modell der visuellen Wahrnehmung komplexer Reizmuster vereinbar, in dem die Repräsentation eines visuellen Musters als Hierarchie von Propositionen zwischen strukturellen Einheiten mit globalen und lokalen Merkmalen nach dem Modell von Norman & Rummelhart (1975) und Palmer (1979) konzipiert ist.

Ein weiteres, hiermit zusammenhängendes Problem ist die Definition des Komplexitätsgrades einer Reizrepräsentanz. Da nach Wolff (1977) die Komplexität der Reizrepräsentanz die Anzahl zu durchwandernder und zu analysierender Ebenen der Repräsentationshierarchie und damit den Zeitbedarf und die Schwierigkeit des Verfügbarmachens reproduktionsrelevanter Information bestimmt, muß die Reizstruktur so beschrieben werden können, daß aus dieser Beschreibung hervorgeht, welche und wieviele Repräsentationsebenen durchlaufen werden müssen, um die reproduktionsrelevante Information verfügbar zu machen. Ohne eine solche Beschreibung wäre die Annahme mehrerer zu durchlaufender Repräsentationsebenen lediglich eine ad-hoc-Interpretation, aber keine Erklärung. Wolff (1977) konnte die Komplexität der von ihm verwendeten Buchstabenzeilen durch das Redundanzkonzept bzw. das Chunkingkonzept erfassen. Wir werden im folgenden Kapitel versuchen, die Chunkingkonzeption auf tonale Strukturen zu übertragen.

Die Aufgabe in einer konkreten Wahrnehmungssituation bestimmt, welche Information reproduktions- bzw. handlungsrelevant ist. Damit ist der Verarbeitungsaufwand in einer Wahrnehmungssituation also sowohl durch die Anzahl zu durchlaufender Repräsentationsebenen determiniert, deren Maximalzahl im Rahmen einer Komplexitätsbeschreibung der Reizstruktur bestimmbar sein muß, als auch durch die Aufgabenstellung, die die niedrigste zu analysierende Repräsentationsebene bestimmt. Die strukturelle Beschreibung des Reizmusters muß also gleichzeitig so beschaffen sein, daß sie die durch die Aufgabe als handlungsrelevant definierte Repräsentationsebene bestimmbar macht. Für den Wahrnehmenden bedeutet das, daß er bereits vor Ausführung der Handlung diejenige Repräsentationsebene, die die handlungsrelevante Information enthält, antizipieren muß; er muß

also über einen H a n d l u n g s p l a n verfügen, der eine eindeutige Beziehung zur wahrzunehmenden Reizstruktur enthält.

Turveys (1973, 1977) Überlegungen zur Handlungssteuerung bei der visuellen Wahrnehmung, die sich auf die theoretischen Arbeiten zur motorischen Koordination und Steuerung von Bernstein (1967) und Greene (1971a,b) beziehen, stellen diese Beziehung zwischen der Struktur der wahrgenommenen Welt und ihrer Repräsentation im Rahmen eines Handlungsplans her:

"... we hypothesize that the properties of the optic array interlace with the prepresentation of an action plan in a systematic fashion: different properties map into different representations ... Suppose that the properties of stimulation relevant to the control of action may themselves be arranged from more complex to more simple; then perhaps we can imagine a natural mapping of these properties onto the unfolding act, a mapping that preserves their order ... With respect to the representation of the action plan in the highest domain, it is not so much that the specification of a subset of large coordinative structures and functions defined on them relates to higher-order properties of the optic array but rather that the description of the plan in action terms and the description of the plan in perceptual terms are dual statements about the same thing" (Turvey, 1977, S. 257 - 258).

Turvey (1977) geht also davon aus, daß der Handlungsplan gleichzeitig unter zwei Aspekten repräsentiert sein muß: Als Bewegungs- oder Handlungsabfolge (funktionaler Aspekt) und als Repräsentation des für die Handlungssteuerung relevanten Ausschnitts der visuellen Welt (struktureller Aspekt). Das Wolff'sche Repräsentationsmodell spezifiziert den strukturellen Aspekt durch die Hierarchie der Repräsentationsebenen und die dadurch festgelegten potentiellen Verarbeitungseinheiten und den funktionalen Aspekt durch den Ein-

schachtelungsprozeß bei der Fokussierung der die Handlungsausführung ermöglichenden Information auf einer Repräsentationsebene.

Turvey (1977) nimmt für die Struktur eines Handlungsplans ebenfalls eine hierarchische Repräsentation an:

"... we might equate the entities in e a c h abstract representation of an act with coordinative structures, remarking that in higher domains an action plan is represented by functions defined over a relatively small number of large and complex coordinative structures and in lower domains by functions over a relatively large number of small and simple coordinative structures" (Turvey, 1977, S. 223).

Insofern als die Repräsentationshierarchie Teil eines Handlungsplans ist, können wir sie als abstrakte strukturelle Repräsentation potentiell handlungsrelevanter Wahrnehmungsinformationen, als implizite Wissensstruktur (Polanyi, 1966) über die Wahrnehmungswelt betrachten und damit als strukturelle Spezifizierung eines kognitiven Schemas (Palmer, 1977; zum Schema-Konzept vgl. auch Bobrow & Norman (1975), Herrmann (1959), Minsky (1975) und Neisser (1967, 1976)). Im aktuellen Vollzug eines Handlungsplans wird jedoch die abstrakte Repräsentationshierarchie durch die im Handlungskontext erforderliche Informationsaufnahme und -verarbeitung auf mehreren Ebenen aktiviert und dient damit der konkreten funktionalen Repräsentation des Wahrgenommenen und der situationsadäquaten Konkretisierung des Handlungsplans.

Das Erkennen der Handlungsrelevanz der auf den verschiedenen Ebenen der Repräsentationshierarchie vorliegenden Informationen ist also unmittelbar vom Wissen des wahrnehmenden Individuums um die Strukturen seiner Wahrnehmungswelt abhängig und von seiner Erfahrung

und Fähigkeit, sie in seine Handlungen einzubeziehen. Den Möglichkeiten des Erkennens musikalischer Strukturen sind im wesentlichen von der vertikalen Differenziertheit des musikalischen Repräsentationssystems ("degree of syntactic articulation", Laske (1977a)) Grenzen gesetzt. Ist die vertikale Ebenengliederung der Repräsentationshierarchie undifferenziert, d.h. besitzt weniger Ebenen, als die Komplexität der musikalischen Strukturen potentiell zuläßt, dann muß eine Strukturerkennungsaufgabe, die ein beim Hörer nicht repräsentiertes Strukturniveau anspricht, von diesem nur unter größeren Schwierigkeiten zu lösen sein, weil er eine handlungsrelevante Information nur über eine zu hohe oder zu niedrige Repräsentationsebene der musikalischen Struktur entnehmen kann.

Im Modell von Wolff (1977) muß angenommen werden, daß zu jedem Zeitpunkt einer Wahrnehmung wenigstens drei vertikal benachbarte Repräsentationsebenen aktiviert sind (Wolff (1977) geht, wenn wir ihn richtig interpretieren, allerdings von nur zwei aktivierten Ebenen aus): 1. Die jeweils gerade fokussierte Ebene, deren Verarbeitungseinheiten im Sinne Navon's (1977) gerade lokal analysiert werden, 2. die unmittelbar übergeordnete Ebene, die die sukzessive Fokussierung der Verarbeitungseinheiten steuert und damit im Sinne Navon's eine globale Analyse der fokussierten Ebene leistet, und 3. die unmittelbar der fokussierten Ebene untergeordnete Ebene, die aufgrund der lokalen Analyse der fokussierten Ebene, die zugleich eine globale Analyse der untergeordneten Ebene ist, eine globale Repräsentation der Konstituenten der Verarbeitungseinheiten auf der fokussierten Ebene erstellt.

Diese dreifache Aktivierung von Repräsentationsebenen ist eigentlich phänomenal unmittelbar einsichtig: Wenn ich das gerade auf meinem Schreibtisch liegende Buch betrachte, dann ist dieses vom Untergrund

der Schreibtischplatte abgehoben, aber der Schreibtisch bleibt phänomenal als übergeordnetes Objekt repräsentiert; gleichzeitig aber ist nicht nur das Buch mit seiner Form, Größe und Farbe lokal repräsentiert, sondern auch die Tatsache, daß der Buchdeckel Schriftzeichen und eine Abbildung trägt, obwohl die Schriftzeichen nicht als Worte einer bestimmten Bedeutung und die Abbildung nicht mit einem bestimmten Inhalt repräsentiert sind, sondern nur global als überhaupt vorhanden in Erscheinung treten.

Diese mindestens dreifache Aktivierung unmittelbar benachbarter Repräsentationsebenen ist funktional schon deshalb notwendig, um einen Übergang von einer zur anderen Repräsentationsebene in b e i d e n Richtungen zu ermöglichen. Wolff (1977) beschreibt den Vorgang der Fokussierung wahrscheinlich nur deshalb in einer Richtung, weil in den von ihm untersuchten Aufgabensituationen (Buchstabenidentifikation) tatsächlich nur ein schrittweiser Übergang von höheren zu niedrigeren Repräsentationsebenen notwendig ist.

Aber selbst beim Lesen ist manchmal ein Wechsel der fokussierten Repräsentationsebene in beiden Richtungen erforderlich, nämlich immer dann, wenn bei schwierigen Wörtern (z.B. einer Fremdsprache) auf das Silben- oder gar Buchstabenniveau gewechselt werden muß und anschließend wieder auf Wortniveau oder gar ein höheres zurückgekehrt wird. Neumann (198o) bezeichnet diesen Wechsel der Repräsentationsebene als F o k u s s i e r u n g s ä n d e r u n g im Gegensatz zum k o n s t a n t e n F o k u s s i e r e n eines Objekts auf einem bestimmten Repräsentationsniveau.

Wenn eine Repräsentationshierarchie aber in beiden Richtungen durchlaufen werden kann, dann ist die funktionale Hierarchie der beteiligten Prozesse als interaktiv (Neumann, 198o) oder als Heterarchie (Turvey, 1973, 1977) aufzufassen:

"From a hierarchical view we might think of perception as an ordered sequence of unidirectional mappings from less abstract to more abstract representations and the differentiation of an intention as the successive mappings of the intention onto a series of progressively l e s s abstract representations. But the argument from coalitional/heterarchical interpretations of organization is that the conversation between abstract representations (domains, knowledge structures) is not one way. A fundamental result in artificial intelligence research on scene analysis is that while it is necessary to construct descriptions in many different domains, a procedure that exploits only unidirectional mappings from lower domain to the next and higher domain is significantly limited in its capability to interpret a scene successfully (Sutherland, 1973). Success in scene interpretation is gradually enhanced by allowing a more flexible strategy in which processing in lower domains can use, as subprocedures, hypotheses generated about structures in higher domains (e.g. Falk, 1972)" (Turvey, 1977, S. 222 - 223; zur hier angesprochenen Entwicklung interaktiver Programme und ihrer Bedeutung für die Wahrnehmungstheorie vgl. Turvey, 1975).

"In one, the contrast is between the hierarchic strategy of a detached higher level dictatorially commanding lower levels and the heterarchic strategy of procedures constructing a representation in a higher domain entering into 'negotiations' with lower domains in order to determine how the higher representation should be stated. In the other, the contrast is between the hierarchic principle of low-level structures unquestioningly responding to high-level instructions and the heterarchical principle of procedures establishing a representation in a lower domain reprocessing higher representations from the perspective of special kinds of knowledge available to the lower domain" (Turvey, 1977, S. 223 - 224).

Das Heterarchieprinzip ist zur Beschreibung der auf die strukturelle Repräsentationshierarchie zugreifenden Prozesse auch von größerer theoretischer Integrationskraft, weil durch die Einführung des Feedbackprinzips für beide Richtungen der Repräsentationshierarchie nicht nur die aktuelle Informationsverarbeitung bei der Wahrnehmung modelliert werden kann, sondern auch die

Änderung und Umorganisation vorhandener Wissensstrukturen aufgrund neuer Erfahrungen, d.h. in dieser Fassung ist das Modell der strukturellen Repräsentationshierarchie und funktionalen Heterarchie sowohl ein Modell der Wahrnehmung als auch ein Modell des Wahrnehmungslernens und der Entstehung von Wissensstrukturen.

Das im Verhältnis zu einem linearen Modell größere Potential einer heterarchischen Systemkonzeption wird durch das von Winograd (1972) entwickelte Computerprogramm zur Analyse von Sprache veranschaulicht. Während ältere Modelle des Verstehens von Sprache (z.B. Johnson, 1965) davon ausgingen, jeden einzelnen Satz erst syntaktisch und dann semantisch zu betrachten und verbleibende Inkonsistenzen anschließend mit Hilfe schlußfolgernden Denkens zu lösen (also durch die Annahme dreier linear verknüpfter Systeme), sieht Winograd's Programm ein simultanes Operieren aller drei Analysesysteme vor, die außerdem noch miteinander in Interaktionen treten. Auf diese Weise ist das System zu bemerkenswerten Leistungen des "Verstehens" seiner (allerdings reduzierten semantischen Welt) fähig: Es beschränkt sich nicht auf das bloße Abrufen gespeicherter Daten und Regeln, sondern zieht Schlüsse aus vorhandenem Wissen und eingegebenen sprachlichen Informationen, verknüpft alte und neue Informationen bei der Beantwortung von Fragen und begründet sogar seine Antworten. Das "Verstehen" beschränkt sich außerdem nicht nur auf isolierte Sätze, wie dies noch bei der Tiefenstrukturanalyse im Modell von Fodor, Bever & Garrett (1976) in Anlehnung an Chomsky's (1965) Syntaxmodell der Fall ist, sondern kann die in vorausgegangenen Sätzen enthaltenen Informationen berücksichtigen.

Wenn heterarchische Modelle die offensichtlichen theoretischen Schwierigkeiten linearer Modelle zu überwinden vermögen, dann könnte ein solches Modell auch den konzeptionellen Rahmen eines kognitiven Modells des Musikhörens liefern, vorausgesetzt es gelingt, Musik als hierarchisch strukturiert zu beschreiben und den Nachweis zu führen, daß sie beim Hören hierarchisch repräsentiert wird.

3.4 ZUSAMMENFASSUNG

Die als ungelöst identifizierten theoretischen Probleme der auditiven Mustererkennung - das der Ganzheitlichkeit der Wahrnehmung und Repräsentation und das der differentiellen Schwierigkeiten der Identifikation von tonalen Strukturen bei verschiedenen Konfigurationen - stellen sich für die Theorien der visuellen Wahrnehmung in gleicher Weise, scheinen uns dort aber einer theoretisch befriedigenden Lösung weit näher zu sein.

Durch zahlreiche Demonstrationen des Einflusses konfigurativer Merkmale auf die Prozesse der visuellen Wahrnehmung, die nicht als konjunktive Verknüpfungen isolierter Features interpretierbar sind, wurde das schwächste Glied des linearen Informationsverarbeitungsansatzes, nämlich die Annahme einer Konstruktion von Wahrnehmungsrepräsentanzen auf der Grundlage unabhängig voneinander analysierter Features, so stark belastet, daß die theoretische Adäquatheit einer Erklärung durch lineare Informationsverarbeitungsmodelle grundsätzlich angezweifelt werden muß: "Sequential arguments are only as strong as their weakest links" (Shaw & Bransford, 1977, S. 12-13).

Eine alternative theoretische Konzeption deutet sich im funktionalen Modell der globalen und lokalen Analyse von Navon (1977) und im Modell der strukturellen Repräsentationshierarchie von Palmer (1977) an. Funktionale und strukturelle Aspekte der kognitiven Verarbeitung integriert das bislang wohl am weitesten ausgearbeitete Modell der strukturellen Repräsentationshierarchie und funktionalen Heterarchie, das zuerst von Wolff (1977)

formuliert und von Neumann (1980) im Zusammenhang mit der Erklärung des Stroop-Interferenzphänomens (Stroop, 1935) in einem umfassenden Modell der kognitiven Handlungssteuerung weiterentwickelt wurde.

Im folgenden werden wir prüfen, inwieweit dieses Modell zur Erklärung der Phänomene der auditiven Mustererkennung, speziell zur Beantwortung der offenen theoretischen Fragen beitragen kann. Auf der Grundlage dieses Ansatzes können wir als vorläufige Hypothese formulieren, daß die Reihenfolge der Repräsentation eines tonalen Musters mit der höchsten strukturellen Repräsentationsebene beginnen sollte und daß die Schwierigkeit der Identifikation von strukturellen Details von der Anzahl zu durchlaufender Repräsentationsebenen abhängen sollte. Das noch zu lösende theoretische Problem besteht in der Definition der strukturellen Komplexität einer Tonfolge, die die Anzahl potentieller Repräsentationsebenen bestimmt. Diesem Problem wenden wir uns nun im folgenden Kapitel zu.

4. AUDITIVES CHUNKING: DIE BILDUNG FUNKTIONALER EINHEITEN DER AUDITIVEN ANALYSE

Die Diskussion von Befunden und theoretischen Ansätzen zur visuellen Wahrnehmung komplexer Reizmuster hat uns weitere Argumente an die Hand gegeben, die, wenn wir sie auf die auditive Wahrnehmung übertragen, eine Revision und Relativierung der Annahme erforderlich machen, daß der einzelne Ton die funktionale Einheit bei der Analyse einer Tonfolge oder Melodie sei. Die Frage, um die es nun geht, ist die nach den konfigurativen Strukturen von Tonfolgen, die für die auditive Analyse die Größe und Anzahl funktionaler Einheiten auf jedem p o t e n t i e l l e n N i v e a u einer Repräsentationshierarchie determinieren. Wir sprechen hier von "potentiellen Niveaus" einer Repräsentationshierarchie, weil wir zunächst einmal von einem idealen Hörer ausgehen, der in der Lage wäre, alle nur denkbaren Korngrößen einer Tonfolge veridikal zu repräsentieren. Die im aktuellen Wahrnehmungsprozeß aktivierbaren Repräsentationsebenen hängen, wie wir bereits im vorausgegangenen Kapitel erwähnten, von der Differenziertheit der kognitiven Schemata ab, die einem Hörer für die Wahrnehmung von Musik zur Verfügung stehen.

Zunächst muß die strukturelle Komplexität einer Tonfolge aufgeklärt werden, indem gezeigt wird, bei welchen physischen Reizkonstellationen innerhalb einer Tonfolge welche Zusammenschlüsse von Tönen zu einem Merkmalsverband bzw. einer funktionalen Einheit entstehen. Empirisches Material, das die Prinzipien der Bildung funktionaler Einheiten demonstriert, liefern die in Kapitel 2 referierten Befunde zur wahrgenommenen Organisation von Ton-

folgen, zur rhythmischen Gliederung einer Melodie und zur Integration und Segmentierung bei Variation der zeitlichen Nähe und/oder der Ähnlichkeit der Tonhöhen.

Wenn es gelingt, die vollständige potentielle Repräsentationshierarchie einer Tonfolge zu beschreiben, d.h. durch Anzahl und Korngröße der Verarbeitungseinheiten auf allen Niveaus zu spezifizieren, dann sollte im Kontext einer bestimmten musikalischen Strukturidentifikationsaufgabe das aufgabenrelevante Repräsentationsniveau, die Anzahl der bei einer Analyse zu durchlaufenden Repräsentationsniveaus und die Anzahl der pro Zeiteinheit zu analysierenden Verarbeitungseinheiten bestimmbar und damit die relative Schwierigkeit und/oder der relative Zeitbedarf für die Durchführung einer solchen Aufgabe vorhersagbar sein.

Zuallererst gilt es, die potentiellen Repräsentationsniveaus beim Hören von Tonfolgen zu identifizieren. Hierbei stoßen wir auf ein Problem, das eine Modifikation bzw. weitere Präzisierung des Rahmenmodells der strukturellen Repräsentationshierarchie und funktionalen Heterarchie (fortan kurz als H i e r a r c h i e m o d e l l bezeichnet) notwendig macht. Im Unterschied zu Wolff (1977) und Neumann (1980), die dieses Modell auf der Grundlage von Versuchen mit tachistoskopischer Darbietung stationärer Reizmuster entwickelt haben, beschäftigen wir uns mit der Wahrnehmung von Reizmustern, die sich erst in der Zeit entwickeln. Während also Informationen über stationäre Reizmuster simultan auf einem globalen Niveau repräsentierbar sind, können selbst globale Informationen über zeitliche Reizmuster nur für gewisse kurze Zeitausschnitte simultan repräsentiert sein. Die Annahme, daß eine Repräsentationshierarchie

immer vom globalsten zum aufgabenrelevanten, detaillierten Niveau durchlaufen werden müsse (Wolff, 1977), ist für zeitliche Reizmuster auf solche Repräsentationsniveaus einzuschränken, deren Informationen auch über die Zeit hinweg simultan verfügbar bleiben.

Simultan verfügbar kann nur solche Information sein, die durch eine Verarbeitungseinheit repräsentiert ist. Ansätze zu Modellen der Bildung von Verarbeitungseinheiten in Tonfolgen finden wir bei Dowling (1973a), Neisser & Hirst (1974) sowie Warren (1974a; Warren & Byrnes, 1975). Die Gemeinsamkeiten dieser Modelle lassen sich unter das Konzept des C h u n k i n g s subsummieren; die jeweiligen Unterschiede aber lassen den Ort des Chunkingmechanismus an verschiedenen Stellen innerhalb der Verarbeitungsstufen linearer Informationsverarbeitungsmodelle vermuten, die sich teilweise mit den drei Chunking-Konzeptionen Miller's (1956a: Chunking auf der Ebene des Kurzzeitgedächtnisses), Johnson's (1973, 1978b: Chunking bei Entnahme aus dem sensorischen Speicher) und Wolff's (1977: v i s u e l l e s C h u n k i n g auf dem Niveau des VIS v o r einer Entnahme) zur Deckung bringen lassen.

In Analogie zum visuellen Chunking bei Wolff (1977) begründen wir dann die funktionale Notwendigkeit eines a u d i t i v e n C h u n k i n g s auf dem Niveau des echoischen Speichers als Mechanismus der Bildung von simultan verfügbaren Verarbeitungseinheiten auf einem mittleren Repräsentationsniveau. Die funktionalen Konsequenzen einer derartigen auditiven Chunking-Konzeption für die Selektion und Identifikation struktureller De-

tails in Tonfolgen werden wir am Beispiel des Blanking-Effekts (Divenyi & Hirsh, 1975) diskutieren.

4.1 BILDUNG VON REPRÄSENTANZEN UND AKTIVIERUNG STRUKTURELLER REPRÄSENTATIONSNIVEAUS BEI DER AUDITIVEN WAHRNEHMUNG

Von manchem großen Komponisten wird anekdotenweise überliefert, er habe eine ganze Sinfonie vollständig im Geiste präsent gehabt und die Musik in einem Zuge zu Papier gebracht. Ungeachtet des Realitätsgehalts solcher Anekdoten kann man aber wohl davon ausgehen, daß der Komponist eine sehr globale, aber simultan verfügbare Repräsentation ("semantic model", Laske, 1977a) zum Ausgangspunkt einer Komposition nimmt, die zwar nicht unbedingt ein vollständiges Musikstück zu umfassen braucht, sich aber im Falle der akustischen Realisation über eine weitaus größere Zeitspanne erstrecken würde, als diejenige, die der Auffassungsspanne eines Hörers entspricht. Dieser Sachverhalt mag auch der Grund dafür sein, warum Musikwahrnehmung häufig als kognitive Umkehrung des kompositorischen Prozesses aufgefaßt wird. Die Logik dieses Umkehrschlusses führt zu der bereits "widerlegten" Hypothese, daß dasjenige, was in der Komposition zuletzt ausgeführt wird (nämlich der einzelne Ton) vom Hörer zuerst repräsentiert werden müsse. Eine theoretische Alternative, die weder dieser Logik zum Opfer fällt noch die Annahme einer so globalen ersten Repräsentation notwendig macht, wie sie der Komponist besessen haben mag, als er die Musik zu Papier brachte, bietet unser Hierarchiemodell.

Spezifizierung des Hierarchiemodells für die Wahrnehmung auditiver Reizmuster

Bei der Übertragung des Hierarchiemodells von Wolff (1977), speziell des funktionalen Aspekts des Aufbaus und Durchlaufens der Repräsentationshierarchie, ergibt sich ein Problem, das in der zeitlichen Erstreckung auditiver Reizmuster begründet ist und sich natürlich auch für alle visuellen Reizmuster, die sich in der Zeit entfalten, in gleicher Weise stellt: Zeitlich strukturierte Reizmuster sind, sofern sie die Auffassungsspanne überschreiten, nicht zu jedem Zeitpunkt des Wahrnehmungsprozesses simultan als Ganzes verfügbar. Simultan verfügbar ist jeweils nur derjenige Ausschnitt, der innerhalb des als psychische Präsenzzeit bezeichneten gleitenden Zeitausschnitts liegt (vgl. Michon, 1978).

"Eine Fokussierung auf einem unteren Repräsentationsniveau setzt die Durchwanderung a l l e r Repräsentationsebenen höherer Ordnung voraus" (Wolff, 1977, S. 425, meine Hervorhebung) kann für zeitliche Reizmuster also nur bedeuten, daß eine Fokussierung auf einem unteren Repräsentationsniveau die Durchwanderung aller bereits zu einem bestimmten Zeitpunkt gebildeten Repräsentationsniveaus höherer Ordnung voraussetzt und nicht aller potentiellen Repräsentationsniveaus, die die strukturelle Gliederung eines Reizmusters als Ganzes zuläßt. Im Falle statischer Reizmuster, wie die von Wolff (1977) verwendeten Buchstabenzeilen, ist die Anzahl potentieller Repräsentationsniveaus mit der Anzahl aktuell gebildeter Niveaus identisch, d.h. in jedem Augenblick des Wahrnehmungsprozesses ist die globalste Repräsentanz, nämlich die des

gesamten Reizmusters, ebenso potentiell fokussierbar, wie die lokalste Repräsentanz, die eines Merkmals eines einzelnen Elements.

Anders dagegen sind die Verhältnisse z.B. bei der Wahrnehmung einer Liedstrophe: Bei Erklingen der ersten Töne brauchen diese zwar keineswegs als individuelle Einheiten repräsentiert zu werden, aber bevor die psychische Präsenzzeit überschritten wird, muß eine globale Repräsentanz dieser ersten Töne erstellt worden sein, damit später überhaupt eine Gesamtrepräsentation einer Liedstrophe möglich ist. Diese erste globale Repräsentanz, die dadurch ausgezeichnet ist, daß ihre zeitliche Erstreckung unterhalb der psychischen Präsenzzeit verbleibt, die aber zugleich globaler als die Repräsentanz der in ihr enthaltenen Elemente ist, wollen wir p r i m ä r e R e p r ä s e n t a n z und das zugehörige strukturelle Niveau als p r i m ä r e s R e p r ä s e n t a t i o n s n i v e a u bezeichnen. Erst nach Abschluß der Herstellung einer ersten primären Repräsentanz ist eine Fokussierung und Analyse der sie konstituierenden Elemente durch sukzessive Einschachtelung auf untergeordneten Strukturebenen möglich, deren Elemente wir zusammenfassend als s u b p r i m ä r e R e p r ä s e n t a n z e n bezeichnen wollen.

Für den weiteren Aufbau der strukturellen Repräsentationshierarchie gibt es nun zwei prinzipiell mögliche Denkmodelle. Das erste Modell geht davon aus, daß das jeweils höchste strukturelle Niveau als erstes repräsentiert wird. Nachdem die erste primäre Repräsentanz hergestellt wurde, kann während der Herstellung der zweiten bereits eine globale Repräsentanz gebildet werden, in der die beiden primären Repräsentanzen als Elemente aufgehen. Damit

liegen dann wenigstens drei strukturelle Repräsentationsniveaus vor: Das der die primären Repräsentanzen konstituierenden Elemente, das der primären Repräsentanzen selbst und das der globalen Repräsentanz beider primären Repräsentanzen. Das erste Modell nimmt nun an, daß dieses globale Niveau weiterhin fokussiert wird und auf das jeweils übergeordnete Niveau gewechselt wird, sobald eine erste fokussierbare Einheit auf diesem Niveau konstituiert wurde. Die funktionale Konsequenz dieses Modells bestünde darin, daß mit zunehmend globaler werdender Repräsentation eines Musikstücks das Erkennen von Details auf niedrigeren Repräsentationsniveaus schwieriger werden sollte, weil bei einer Fokussierung einer sehr niedrigen Ebene zunehmend mehr Repräsentationsniveaus durchlaufen und analysiert werden müßten. Es sollte daher z.B. schwieriger sein, einen Quartsextakkord vor einer Solokadenz als solchen zu identifizieren, je später dieser im Verlauf eines Klavierkonzerts auftritt. Diese Konsequenz widerspricht aber jeder musikalischen Hörerfahrung.

Das Alternativmodell dagegen geht davon aus, daß der Hörer auf dem Niveau der primären Repräsentanzen verbleibt und dieses fokussiert. Während weitere primäre Repräsentanzen gebildet werden, können gleichzeitig globalere Repräsentanzen aufgebaut werden, diese dann kurzzeitig fokussiert und zu noch globaleren Repräsentanzen verschmolzen werden, aber zwischenzeitlich kehrt die Fokussierung immer wieder auf das Niveau der primären Repräsentanzen zurück.

Dieses Modell wird der Phänomenologie musikalischer Hörerfah-

rungen gerecht. Die Korngröße der strukturellen Repräsentation fluktuiert zwar, es gibt aber ein Repräsentationsniveau, das überwiegend fokussiert bleibt und zu jedem Zeitpunkt eine Fokussierung niedrigerer Ebenen erlaubt, deren Genauigkeit und Schwierigkeit annähernd konstant bleibt. Die Genauigkeit und Schwierigkeit einer Fokussierung niedrigerer Ebenen ist nur dann größer, wenn gerade eine Repräsentationsebene fokussiert wird, die über der der primären Repräsentanzen liegt, und gleichzeitig eine Handlung notwendig wird, die eine Fokussierung einer Repräsentationsebene weit unterhalb der der primären Repräsentanzen erforderlich macht, wie dies z.B. bei der Identifikation der Position eines einer Melodie überlagerten Klickgeräuschs der Fall ist. Die Lokalisation einer Klickposition sollte also nicht, wie dies das uns unwahrscheinlich erscheinende erste Modell vorhersagen würde, an späteren Positionen einer Melodie schwieriger als an früheren sein, sondern an allen Positionen, an denen das primäre Repräsentationsniveau fokussiert wird, gleich schwierig sein und nur an denjenigen Positionen schwieriger werden, an denen gerade ein darüberliegendes Niveau fokussiert werden muß. Wir werden auf diese Konsequenz des funktionalen Heterarchiemodells des Aufbaus einer Repräsentationsstruktur beim Hören von Musik in Teil II der Arbeit noch im Detail zurückkommen.

Die hier skizzierte sukzessive Herstellung einer Repräsentationsstruktur zeitlich strukturierter Reizfolgen illustriert noch einmal Abb. 1. Die vertikale Ebenengliederung beginnt auf dem untersten Niveau bei den Elementarrepräsentanzen, die im Falle von Musik durch einzelne Töne gebildet werden. Diese

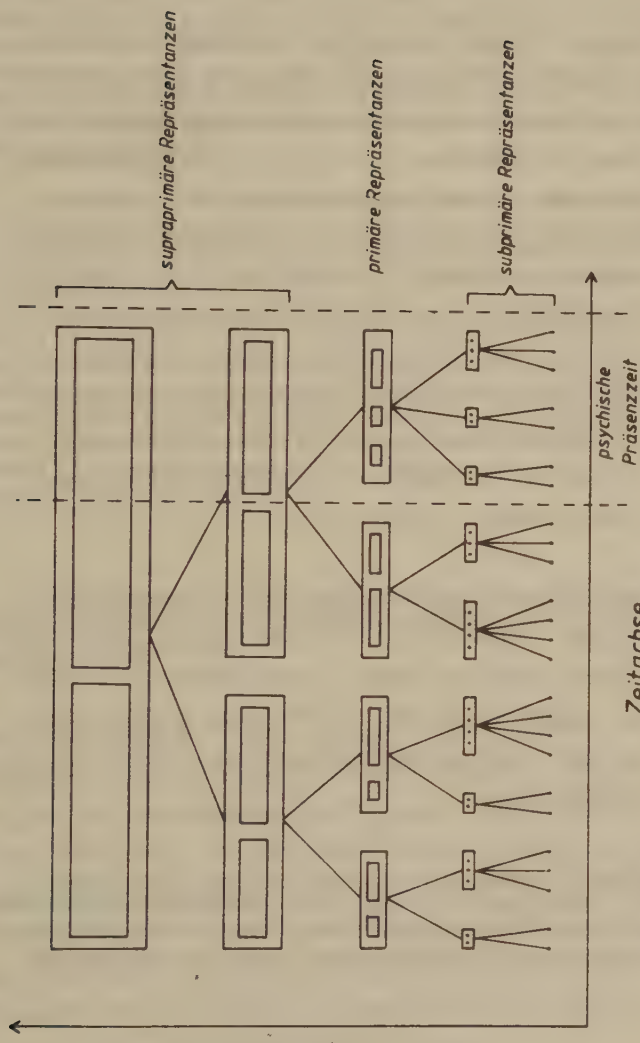


Abb. 1: Repräsentationsniveaus innerhalb und außerhalb der psychischen Präsenzzeit

und die bis zum primären Repräsentationsniveau reichenden Ebenen werden von uns als s u b p r i m ä r e , die dem primären Repräsentationsniveau übergeordneten Ebenen als s u p r a - p r i m ä r e R e p r ä s e n t a n z e n bezeichnet.

Welches strukturelle Niveau zur primären Repräsentationsebene wird, ist im Modell nicht festgelegt. Festgelegt ist lediglich die maximale zeitliche Erstreckung einer primären Repräsentanz durch die psychische Präsenzzeit, die sich in der Größenordnung von wenigen Sekunden bewegen dürfte (Fraisse, 1978; Michon, 1978). So kann es von den individuellen Fähigkeiten und Hörerfahrungen eines Hörers abhängen, welche Ebene er als primäre Repräsentationsebene wählt. Wir wollen an dieser Stelle zumindest erwähnen, daß es wahrscheinlich ist, daß hochmusikalische Hörer im Vergleich zu musikalischen Laien eine höhere Repräsentationsebene wählen, da sie wahrscheinlich über differenziertere kognitive Schemata für globalere Strukturaspekte von Musik verfügen (Näheres hierzu siehe Teil II).

Bei diesem Modell handelt es sich weder um ein reines "bottom-up"-Modell, wie bei Jones (1974, 1978) oder Restle (197o, 1974), noch um ein reines "top-down"-Modell wie bei Navon (1977) oder Palmer (1977), sondern um ein Modell, das beide Verarbeitungsrichtungen auf der Grundlage von Repräsentanzen einer mittleren Korngröße berücksichtigt.

Ein vergleichbares Modell der visuellen Analyse haben Kinchla & Wolfe (1979) formuliert. Sie führten in Anlehnung an die Untersuchung von Navon (1977) ein Experiment durch, in dem die Vpn durch möglichst schnelles Drücken einer Reaktionstaste angeben

mußten, ob ein ihnen vorher als Zielreiz angegebener Buchstabe in einem Display vorkommt oder nicht, das aus einem großen Buchstaben, der aus lauter kleinen Buchstaben zusammengesetzt ist, bestand. Im Unterschied zu Navon (1977) variierten sie auch die Größe des Displays zwischen ca. 5° und 25°. Es zeigte sich, daß entsprechend den Ergebnissen von Navon eine Reaktion auf den großen Buchstaben schneller als auf die kleineren abgegeben werden konnte, wenn der Winkel 9° nicht überstieg. Bei grösseren Winkeln dagegen kehrte sich das Befundmuster um: Die kleineren Buchstaben konnten dann schneller als die großen identifiziert werden.

Die Autoren interpretieren dieses Ergebnis als Zeichen dafür, daß mittlere Strukturniveaus eines Reizmusters vor höheren und niedrigeren verarbeitet werden, wenn die höheren Niveaus eine optimale Größe im visuellen Feld überschreiten. Wenn wir die Größe eines visuellen Reizmusters als funktional vergleichbar der zeitlichen Erstreckung eines auditiven Reizmusters betrachten, dann sollten höhere strukturelle Niveaus eines auditiven Reizmusters später als mittlere Niveaus repräsentiert werden, wenn die höheren Niveaus eine optimale zeitliche Erstreckung überschreiten, wie dies für musikalische Strukturniveaus gilt, deren zeitliche Erstreckung die psychische Präsenzzeit übersteigt.

Funktionale Bedeutung supraprimärer Repräsentationsniveaus beim Musikhören

Supraprimäre Repräsentanzen enthalten die zu jedem Zeitpunkt des Hörvorgangs simultan verfügbaren konfigurativen Informationen über die bis zum gegenwärtigen Zeitpunkt wahrgenommene musikalische Struktur und stellen damit eine in nicht-zeitlicher Form verfügbare Repräsentation sich über die Zeit erstreckender Abläufe dar. So wie in Wolff's (1977) Hierarchiemodell eine Verarbeitungseinheit dadurch konstituiert wird, daß sie die in ihr enthaltene Information über die Reizstruktur simultan verfügbar

macht, indem sie sie an einem Ort innerhalb der horizontalen Gliederung eines Niveaus repräsentiert, so wird die gegenwärtige und/oder vergangene Reizinformation auf supraprimären Niveaus der Repräsentationshierarchie dadurch simultan verfügbar, daß sie als zu einem gemeinsamen Zeitintervall gehörig repräsentiert ist.

Auf den supraprimären Repräsentationsniveaus werden Gegenwart und Vergangenheit aufeinander bezogen. Gerade fokussierte musikalische Abläufe geringer Korngröße (melodische Figuren, harmonische Wendungen, rhythmische Verschiebungen u.dgl.) erscheinen phänomenal auf dem Hintergrund globaler Repräsentationen der Gesamtheit des bis zu diesem Zeitpunkt Gehörten (immer einschränkend vorausgesetzt, der Hörer ist in der Lage, derart globale Repräsentationen aufzubauen). Auf diese Weise können strukturelle Beziehungen zwischen gegenwärtigen und vergangenen Abläufen erkannt werden, wenn ein der zeitlichen Reichweite dieser Beziehungen entsprechendes Repräsentationsniveau fokussiert wird, dessen Korngröße so gewählt wird, daß gegenwärtiger und vergangener musikalischer Ablauf sich auf demselben horizontalen Niveau bewegen und dort fokussierbare Verarbeitungseinheiten darstellen.

Das Erkennen detaillierter Unterschiede zwischen einem Thema und seiner Variation in einem Rondo ist hierfür ein Beispiel. Die zeitlichen Abstände zwischen der Vorstellung des Themas und nachfolgenden Variationen können bis zu mehreren Minuten betragen. Ungeachtet der eventuell vorhandenen Tempounterschiede und Verschiedenheit der Taktzahlen können für einen

derartigen Vergleichsprozeß Thema und Variation als Verarbeitungseinheit bereits relativ großen Korns (es handelt sich hierbei immerhin um Einheiten in der Größenordnung von ca. acht Takten) repräsentiert werden. Je nach zeitlichem Abstand und Komplexität der dazwischen liegenden musikalischen Entwicklung kann die Beziehung dieser Einheiten nur über ein noch weitaus globaleres Repräsentationsniveau gestiftet werden, von dem aus durch den von Wolff (1977) beschriebenen sukzessiven Einschachtelungsprozeß Thema und Variation fokussierbar werden und durch weitere Einschachtelung die sie unterscheidenden Details. Ein musikalischer Vergleich ist also immer nur durch Fokussierung übergeordneter Repräsentationsniveaus möglich ("bottom-up"-Anteil) mit einer anschließenden Fokussierung der für den Vergleich relevanten Detailinformationen ("top-down"-Anteil).

Ein Spezialfall der Beziehungsbildung beim strukturellen Hören von Musik ist die r e t r o s p e k t i v e U m d e u t u n g . Hierbei handelt es sich um die nachträgliche Umdeutung einer mehrdeutigen strukturellen Beziehung aufgrund der Informationen des nachfolgenden Kontextes, wie dies manchmal auch bei der semantischen Interpretation syntaktisch mehrdeutiger Sätze notwendig ist. Charakteristischerweise wird dabei die Grenze zwischen zwei strukturellen Einheiten bzw. deren Repräsentation verschoben. Als Beispiel nehmen wir eine unbegleitete Melodie, deren Töne in den ersten zwei Takten harmonisch eindeutig Tonikafunktion besitzen und denen dann ein Takt mit dem Ton der fünften Stufe folgt. Da Taktgruppierungen in einer musikalischen Periode vom Idealfall der Zweiergruppierung geprägt sind und auf übergeordneten Strukturniveaus sich grad-

zahlige Vielfache von Zweiergruppen zusammenschließen, liegt es für einen Hörer, dessen kognitives Schema eine solche Zweiergruppierung enthält, zunächst nahe, die ersten beiden Tonikatakte als eine strukturelle Einheit zu repräsentieren und den Takt mit der fünften Tonstufe als Beginn einer weiteren Zweiergruppe, also als neue Einheit zu betrachten. Diese Einheitenbildung kann nun bei Wahrnehmung des vierten Taktes, der eindeutig eine Dominantenfunktion besitzen soll, beibehalten werden, denn der dritte Takt ist im nachhinein nun harmonisch ebenfalls als zur Dominante gehörig interpretierbar. Der fünfte und sechste Takt, der harmonisch weiterhin eindeutig Dominantencharakter besitzen soll und mit einer langen Note mit anschliessender Pause schließt, läßt nun die erste Einheitenbildung unplausibel erscheinen, denn offenbar handelt es sich nicht um eine Zweiergliederung der Takte der Melodie, sondern um einen Zusammenschluß von jeweils drei Takten, wobei die ersten drei Takte Tonikafunktion besitzen (einschließlich des dritten Taktes mit der fünften Tonstufe, die ja auch Tonikafunktion besitzen kann) und die folgenden drei Dominantfunktion. Die zuerst vorgenommene Repräsentation in Zweiergruppen muß aufgrund des nachfolgenden Kontextes zu einer Repräsentation von Einheiten aus Dreiergruppen umorganisiert werden. Diese Umorganisation ist nur aufgrund einer globalen Repräsentation aller sechs Takte auf einer supraprimären Repräsentationsebene und Fokussierung einzelner Takte mit anschließendem Neuaufbau einer übergeordneten Repräsentation möglich.

Die funktionale Bedeutung supraprimärer Repräsentationsniveaus erstreckt sich aber nicht nur auf die gleichzeitige Repräsentation vergangener und gegenwärtiger musikalischer Ereignisse,

sondern auch auf die Antizipation zukünftiger musikalischer Entwicklungen auf der Basis gegenwärtiger Repräsentationen. Nach Neisser sind Wahrnehmungsschemata "... plans for finding out about objects and events, for obtaining more information to fill in the format" (Neisser, 1976, S. 55), d.h. als Plan müssen sie den jeweils aktuellen Wahrnehmungen antizipatorisch vorauseilen. Aber diese antizipatorische Bereitstellung geeigneter "Formate" zur selektiven Aufnahme von Informationen muß ja selbst wiederum durch die aktuellen Wahrnehmungen und die resultierenden Repräsentationen gesteuert werden. Eine Antizipation einzelner musikalischer Ereignisse geringer Korngröße ist, wie die musikalische Hörerfahrung häufig bestätigt, nur äußerst selten möglich. Antizipatorische Funktionen können daher wohl nur globale Repräsentationen, wie die supraprimären, übernehmen.

Musikalische Erwartungen stellen also keine Vorwegnahmen einzelner, konkreter musikalischer Ereignisse dar, sondern sind Extrapolationen von auf supraprimären Repräsentationsebenen fokussierten, konfigurativen Informationen auf zukünftige Repräsentationen gleicher Korngröße. Eine Erwartung besteht darin, daß zukünftige musikalische Ereignisse als auf der fokussierten Ebene repräsentierbar und auf der unmittelbar übergeordneten Ebene als Realisation desselben Typs von Relationen oder Konfigurationen, wie die zwischen den Einheiten der fokussierten Ebene, betrachtet werden. Wenn wir ein Musikstück mit deutlich voneinander abgehobenen achttaktigen Perioden hören, dann werden wir dieselbe Gliederungsstruktur auch für zukünftige Entwicklungen des Musikstücks erwarten.

Die hier dargestellten theoretischen Ausführungen sind zugegebenermaßen noch recht spekulativ. Wir haben mit ihnen bereits den theoretischen Rahmen überschritten, der für die Thematik von Teil I von unmittelbarer Bedeutung ist. Doch war es bereits an dieser Stelle notwendig, die theoretischen Spezifikationen des Rahmenmodells von Wolff (1977) für die auditive Wahrnehmung einzuführen, um die zentrale Thematik von Teil I gegenüber der von Teil II deutlicher und theoretisch begründet abgrenzen zu können. In Teil I werden wir uns weiterhin lediglich mit der Wahrnehmung und Repräsentation so kurzer Tonfolgen beschäftigen, deren zeitliche Erstreckung die Zeitspanne der psychischen Präsenzzeit nicht überschreitet. Die im Mittelpunkt der weiteren theoretischen Diskussionen stehenden Repräsentationsniveaus sind also das primäre und die subprimären Repräsentationsniveaus. Im folgenden geht es darum zu begründen, daß die Annahmen über die Reihenfolge der Bildung von globalen und lokalen Repräsentationsniveaus relativ geringer Korngröße für die Wahrnehmung kurzer Tonfolgen zutreffen, um auf dieser Grundlage die in unserer Diskussion der Phänomene der auditiven Mustererkennung formulierten Probleme der Erklärung der Ganzheitlichkeit der Wahrnehmung von Tonfolgen und der differentiellen Schwierigkeiten bei der Identifikation ihrer Strukturen einer Erklärung im Rahmen des Hierarchiemodells näherzubringen.

4.2 VERARBEITUNGSEINHEITEN INNERHALB DER PSYCHISCHEN PRÄSENZZEIT

Welche musikalischen Einheiten können als funktionale Einheiten auf dem primären Niveau repräsentiert sein? Hierbei kann es sich ja nur um solche Einheiten handeln, deren zeitliche Erstreckung

die minimale Zeitspanne der simultanen Verfügbarkeit ihrer Elemente, also die psychische Präsenzzeit, nicht überschreitet. Fraisse (1978) gibt die maximale zeitliche Erstreckung der psychischen Präsenzzeit bei mehr als zwei Tönen mit ca. fünf Sekunden an und stellt gleichzeitig fest, daß selbst im langsamsten Adagio kein Takt länger als fünf Sekunden dauert, d.h. also: Einzelne Takte können keinem höheren Niveau als dem der primären Repräsentanzen zugeordnet werden, sie können wohl aber auf subprimären Niveaus repräsentiert sein, wenn bei hohen Tempi zwei oder mehr vollständige Takte die psychische Präsenzzeit nicht überschreiten.

Verallgemeinern wir diesen Sachverhalt, dann kommen wir notwendigerweise zu der Hypothese, daß mit zunehmender Darbietungsgeschwindigkeit eines zeitlich erstreckten Reizmusters immer mehr strukturelle Niveaus den funktionalen Status einer subprimären Repräsentation einnehmen, wodurch gleichzeitig die Anzahl potentieller supraprimärer Repräsentationsniveaus abnimmt.

Für die Mechanismen der Strukturidentifikation ergibt sich hieraus die Konsequenz, daß die Identifikation bei schnellen Tempi mit einer globaleren Repräsentation beginnt und zum Erkennen von strukturellen Details auf einem niedrigen Niveau (z.B. bei Intervallrichtungsanalyse) mehr Repräsentationsniveaus durchlaufen muß als bei langsamen Tempi, es sei denn, es ließe sich eine niedrigere Ebene konstant fokussieren. Eine Kompensation des erhöhten Verarbeitungsaufwands bei schnellen Tempi durch direkte Fokussierung niedriger Repräsentanzen ist aber im Grenzfall der Merkmalsintegration sukzessiver Töne bei SOAs von weniger als

6o bis 8o msec nicht möglich, wie das Beispiel des Glissandos zeigt, das wir als kontinuierliche Tonhöhenänderung wahrnehmen, also primär als Ganzes repräsentieren, bevor eine eventuelle Detailanalyse in nachträglicher analytischer Betrachtung möglich ist.

Zudem ist eine Fokussierung niedrigerer Repräsentanzen kaum eine erfolgversprechende Kompensationsstrategie, da mit steigenden Tempi auf jeder Repräsentationsebene die Zahl der pro Zeiteinheit zu verarbeitenden Einheiten zunimmt und um so größer ist, je niedriger das fokussierte Niveau ist. Aus dieser Sicht kann man die Wahl eines höheren Repräsentationsniveaus auch dadurch funktional begründen, daß man annimmt, das kognitive System unterliege einer Begrenzung bezüglich der Anzahl der pro Zeiteinheit zu verarbeitenden Einheiten und schütze sich durch Wahl globalerer Einheiten vor einer Überlastung, die notwendigerweise dazu führen müßte, daß einige Einheiten nur auf Kosten anderer, die unanalysiert bleiben müssen, weiterverarbeitet werden können (Argumente für eine Begrenzung der auditiven Analysegeschwindigkeit lassen sich aus Befunden zur rückwirkenden Maskierung (Massaro, 1976a) und aus Untersuchungen zur Genauigkeit des Zählens von Tönen bei unterschiedlichen SOAs (Massaro, 1976b) ableiten). In diesem Zusammenhang ist die Wahl eines globaleren Repräsentationsniveaus ein adaptiver Mechanismus des kognitiven Systems, ein Optimum an Informationen aus einem Reizmuster zu gewinnen, ohne die Möglichkeit einer gesteuerten Selektion aufzugeben, wie es bei Fokussierung eines zu niedrigen Niveaus bei zu hohem Tempo der Fall wäre.

Eine Erklärung des Einflusses verschiedener Darbietungsgeschwindigkeiten auf die Identifikation von Tonfolgen durch die Annahme unterschiedlicher Repräsentationsformen haben auch Neisser & Hirst (1974) und Warren (1974a) vorgeschlagen. Beide Modelle gehen allerdings von nur zwei qualitativ unterschiedlichen Repräsentationsformen aus: eine Gesamtrepräsentation des akustischen Reizmusters ("analogs" bei Neisser & Hirst (1974); "temporal compounds" bei Warren (1974a)), auf die sich die Identifikation bei hohen Darbietungsgeschwindigkeiten stützt, und eine Repräsentation einzelner Reizmusterelemente ("strings" bei beiden Modellen), die Gegenstand der Identifikation bei geringen Darbietungsgeschwindigkeiten sind. Insbesondere Warren's (1974a) Konzept der "temporal compounds", denen überadditive Eigenschaften zugesprochen werden, hat starke Ähnlichkeit mit dem "chunking"-Konzept bei Dowling (1973a): "temporal compounds" werden, wie auch Chunks, zunächst als ganze Verarbeitungseinheiten repräsentiert, die erst durch zusätzliche Verarbeitungsschritte in ihre konstitutiven Elemente analytisch zerlegt werden können. Unklar bleibt aber sowohl bei Warren (1974a) als auch bei Dowling (1973a) der Ort des Chunkings innerhalb der als lineare Abfolge konzipierten Verarbeitungsschritte.

Bezüglich der Bildung von Verarbeitungseinheiten durch Chunking lassen sich drei Modelle unterscheiden: 1. Chunking auf dem Niveau des Kurzzeitgedächtnisses (Miller, 1956a,b), 2. Chunking als Kodierungsschritt zwischen sensorischer Speicherung und Einlesen ins Kurzzeitgedächtnis (Johnson, 1973, 1978b) und 3. Chunking als "Primärkodierung" vor einer sensorischen Speicherung (Wolff, 1977).

Miller's (1956a,b) Chunking-Konzeption, der zunächst auch Johnson (1965, 1969, 197o, 1972) und Dowling (1973a) folgten, beschreibt einen Mechanismus der Kodierung von Einheiten, die sequentiell dem sensorischen Speicher entnommen und ins Kurzzeitgedächtnis eingelesen werden, wo sie aufgrund zeitlicher Nähe und/oder rhythmischer Gruppierung zu Einheiten höherer Ordnung r e k o d i e r t werden. Diese Chunking-Konzeption dient der funktionalen Erklärung einer einheitsspezifischen statt informationsspezifischen Beschreibung des Kurzzeitgedächtnisses (Bower, 197o, 1972; Bower & Springston, 197o; Bower & Winzenz, 1969, Murdock, 1961; Pollack & Johnson, 1965; Posner, 1963; Winzenz, 1972).

Johnson's (1973, 1978b) Chunking-Konzeption unterscheidet sich von der Miller's (1956a,b) darin, daß der Ort des Chunkingprozesses auf eine Verarbeitungsstufe vor der einer kurzzeitigen Speicherung verlagert wird. Der Anlaß hierfür waren die Befunde eines Experiments (Johnson, 1973), in dem sich herausstellte, daß die Identifikation einer deutlich durch Zwischenräume markierten Einheit von drei Buchstaben schneller erfolgen konnte als die Identifikation der die Einheit bildenden Buchstaben. Dieser Befund läßt sich nur dahingehend interpretieren, daß Einheiten bereits während der Wahrnehmungsprozesse gebildet werden und als Ganze ins Kurzzeitgedächtnis gelangen.

Auf dieser Grundlage gelangte Johnson (1978b) zu einem vierstufigen Modell der visuellen Identifikation. Auf einer ersten Stufe wird ein Reiz als physische Information in Einzelkomponenten repräsentiert. Auf der zweiten Stufe findet während der

Informationsentnahme am VIS ein "unitization process" statt, der die Einzelkomponenten durch Kontakt mit Gedächtniskodes (phonetische Kodes auf Silben- oder Wortniveau) zu Verarbeitungseinheiten verschmelzen läßt, die als Einheit einem Gedächtniskode zuzuordnen sind. Findet sich kein passender Gedächtnisinhalt (wie z.B. bei sinnlosen, nicht aussprechbaren Silben), werden dem VIS die Einzelkomponenten sequentiell entnommen. Der auf der zweiten Stufe erstellte Kode in Form von Chunks oder Einzelkomponenten gelangt auf der dritten Stufe in den Arbeitsspeicher, wo eine semantische Interpretation stattfindet. Auf der vierten Stufe schließlich werden zu vergleichende Kodes (z.B. same/different-Vergleich der Silben HUT und HET) in eine den Vergleich ermöglichende Form gebracht (in diesem Beispiel müßte also der Kode /HUT/ fraktioniert werden zu /H/U/T/, um mit den drei Kodes /H/E/T/ verglichen werden zu können).

In Wolff's (1977) Modell der Redundanzausnutzung bei tachistoskopisch dargebotenen Buchstabenfolgen wird der Ort des Chunkings um eine weitere Stufe gegenüber dem Modell von Johnson (1978b) vorverlagert:

"Im Rahmen unseres Modells lassen sich die Chunkinghypothese und die Hypothese der Abtastbeschleunigung bei redundanten Reizen nur dadurch zu einer geschlossenen Theorie der Redundanzausnutzung integrieren, daß man den Chunking-Prozeß als eine Eigenschaft der Primärkodierung auffaßt und annimmt, daß die Chunks bereits in der primären Reizrepräsentanz vorliegen. Die dynamische Eigenschaft der Abtastung, bei redundanten Reizen beschleunigt zu operieren, läßt sich strukturell damit begründen, daß die Chunks redundanter Reize den Elementen zufälliger Buchstabenfolgen als Abtasteinheiten funktional äquivalent

sind... Nach unserer Theorie ist ferner die Erweiterung der Gedächtnisspanne bei redundanten Reizen dadurch zu erklären, daß die redundanten Reize bereits als eine Folge von Chunks in das Kurzzeitgedächtnis gelangen. Ein der Abtastung nachfolgender Chunking-Prozeß im Kurzzeitgedächtnis erübrigt sich demnach" (Wolff, 1977, S. 135-136).

Wolff's (1977) Konzeption eines "visuellen Chunking" besitzt gegenüber der von Johnson (1978b) den theoretischen Vorteil, ohne eine Gedächtnissuche nach phonetischen Einheiten, die die Elementrepräsentation im VIS zu einer Chunkrepräsentation verschmelzen lassen könnten, auszukommen. Die Gedächtnissuche muß ja außerdem irgendwie gesteuert sein; wenn diese Steuerung aber aufgrund der Elementrepräsentation im VIS vollbracht würde, dann müßten die Elementrepräsentanzen erst einmal sequentiell dem VIS entnommen werden, wie dies Johnson (1978b) ja auch sieht, aber damit auch als solche identifiziert und bewußt repräsentiert sein, bevor sie zu Chunks verschmelzen. Dies aber ist offenbar nicht der Fall, denn sonst wären die Befunde von Johnson (1973) nicht zu erklären.

4.3 AUDITIVES CHUNKING

Wenn wir in Anlehnung an die Chunking-Konzeption von Wolff (1977) annehmen, daß dem echoischen Speicher nicht die Information über einzelne Töne, sondern Chunks entnommen werden, deren Organisation von der Konfiguration der Intervalle, der Pausen zwischen den Tönen, der rhythmischen Akzentverteilung und der physischen Ähnlichkeit der Töne abhängt, dann ist 1. erklärbar, daß schnelle, kurze Tonfolgen ganzheitlich repräsentiert werden,

nämlich als ein Chunk, und 2. kann erklärt werden, warum und in welcher Weise die Schwierigkeit der Identifikation der Struktur von Tonfolgen von der Organisation der Tonfolgen abhängt. Bevor wir diese Erklärungen weiter ausführen, müssen wir zunächst überprüfen, ob die Konzeption eines a u d i t i v e n C h u n k i n g s mit Befunden zur echoischen Speicherung und auditiven Analyse vereinbar sind.

Empirische Belege für ein auditives Chunking

Es gibt mindestens zwei Bedingungen, unter denen experimentell zwischen einer Chunking-Konzeption im Sinne von Wolff (1977) und der von Miller (1956a) für die Identifikation tonaler Strukturen entschieden werden kann: 1. Chunking als Rekodierung auf dem Niveau des Kurzzeitgedächtnisses (Miller, 1956a) sollte in gewissen zeitlichen Grenzen um so effektiver sein, je mehr Zeit für die einzelnen Rekodierungen zur Verfügung steht, da dann die dem echoischen Speicher entnommene Information durch Rehearsal stabilisiert werden könnte. 2. Rückwirkende Maskierung, die nach der gängigen Auffassung von den Funktionscharakteristika des echoischen Speichers die Entnahme der Tonhöheninformation beeinträchtigt (Massaro, 197o, 1972), sollte sich bei Zutreffen der Wolff'schen Chunking-Konzeption, wonach die Chunks bereits im sensorischen Speicher vorliegen, nur auf Töne beziehen, die zu einer gemeinsamen Gruppe gehören, nicht dagegen auf Töne, die zu verschiedenen Gruppen organisiert wurden, denn nur dann, wenn die Informationen im echoischen Speicher bereits in organisierter Form vorliegen, ist eine von der Ähnlichkeit physischer Merkmale der Test- und Maskiertöne abhängige, s e l e k t i v e

S t ö r u n g der Informationsentnahme aus dem echoischen Speicher denkbar.

Die Hypothese einer bereits im echoischen Speicher vorliegenden Organisation tonaler Informationen nach physischen Merkmalen wird durch die Befunde von Idson & Massaro (1976) unterstützt. In einer Serie von Maskierungsexperimenten versuchten die Autoren, den funktionalen Mechanismen des Streamingeffekts auf die Spur zu kommen. Dabei verwandten sie das klassische Paradigma der rückwirkenden Maskierung.

Der typische Befund solcher Untersuchungen bei rückwirkender Maskierung eines Testtones durch einen Maskierton besteht darin, daß die Präsenz eines Maskiertons mit der Identifikation des vorausgehenden Testtons interferiert, wenn das ISI[1)] kleiner als 25o msec ist (Massaro, 197o). Den funktionalen Zusammenhang zwischen ISI und Identifikationsrate beschreibt eine negativ beschleunigte Funktion, die bis ca. 25o msec ansteigt und von da ab asymptotisch zur Zeitachse (ISI) verläuft. Dieser funktionale Zusammenhang wird als Beleg für die Existenz eines "preperceptual auditory store" (PAS) gewertet (Massaro, 197o, 1972, 1976a; Massaro & Idson, 1976, 1977, 1978; mit einer auditiven Teilberichtstechnik arbeiteten: Darwin, Turvey & Crowder, 1972; Moray, Bates & Barnett, 1965; Treisman & Rostron, 1972). Aus dem PAS kann solange kontinuierlich Information entnommen werden, bis ein weiterer Ton mit der PAS-Repräsentanz des ersten interferiert.

Idson & Massaro (1976) boten ihren Vpn im Unterschied zum Standardmaskierungsexperiment drei verschiedene Testtöne dar, denen jeweils ein Maskierton folgte. Die Vpn hatten die Aufgabe zu entscheiden, welche der sechs möglichen Permutationen der drei Test-

---

1) "interstimulus interval": Zeitintervall zwischen Darbietungsende eines Reizes und Darbietungsbeginn des folgenden Reizes

töne ihnen vorgespielt wurde. Außer einer Variation des ISIs zwischen Test- und Maskiertönen wurde die Tonhöhe der Maskiertöne relativ zu der der Testtöne variiert und eine Kontrollbedingung eingeführt, in der die Testtöne ohne Maskiertöne dargeboten wurden.

Die Befunde zeigen deutlich, daß bei Maskierung durch Töne aus demselben Frequenzbereich, aus dem auch die Testtöne stammen, der Prozentsatz richtig identifizierter Tonfolgen signifikant niedriger liegt als bei Maskierung durch einen um eine Oktave höheren oder niedrigeren Ton. Bei Verwendung von Maskiertönen außerhalb der Oktave der Testtöne ist die Identifikationsrate nicht verschieden von der in der Kontrollbedingung ohne Maskierung. Die typische negativ beschleunigte Maskierfunktion zeigt sich nur in der Bedingung, in der die Maskiertöne zwischen dem tiefsten und höchsten Testton liegen. Die Identifikationsrate bei Maskiertönen außerhalb der Oktave der Testtöne bleibt unabhängig vom ISI mit ca. 8o % konstant. Im Standardmaskierungsversuch dagegen ist unabhängig von der Größe des Tonhöhenintervalls zwischen Test- und Maskierton immer eine typische Maskierfunktion zu beobachten (Idson & Massaro, 1976, Experiment 3).

Eine Maskierwirkung von Tönen bei der Identifikation der Struktur einer Tonfolge ist also von der Ähnlichkeit der Tonhöhen von Test- und Maskiertönen abhängig. Wenn man voraussetzt, daß die Maskierwirkung in einer Beeinträchtigung der Entnahme von PAS-Information besteht, dann kann dieses Befundmuster nicht damit erklärt werden, daß Informationen über die individuellen

Tonhöhen dem PAS entnommen und anschließend im Kurzzeitgedächtnis organisiert werden. Vielmehr muß die Organisation bereits vorliegen, wenn die Maskiertöne eintreffen, d.h. die gesamte dem PAS entnehmbare Information muß bereits in einer nach physischen Merkmalen organisierten Form vorliegen, wenn ein Maskierton eintrifft, der dann selbst diesem von Wolff (1977) als Primärkodierung bezeichneten Organisationsprozeß unterworfen und nach physischer Ähnlichkeit mit den bereits vorhandenen Tonhöheninformationen gruppiert wird. Bei großer physischer Ähnlichkeit von Maskier- und Testtönen wird dann nur ein Chunk gebildet, aus dem die Testtöne erst durch selektive Fokussierung einzeln entnommen werden müssen, um eine richtige Musteridentifikation zu ermöglichen, während bei geringer Ähnlichkeit zwei Chunks vorliegen, das der Testtöne und das der Maskiertöne, so daß die aufgabenrelevante Information ohne einen zusätzlichen Analyseschritt der Chunkrepräsentanz der Testtöne entnommen werden kann.

Erklärung differentieller Schwierigkeiten bei der auditiven Mustererkennung durch das Modell des auditiven Chunkings und das der Repräsentationshierarchie

Divenyi & Hirsh (1975) griffen noch einmal einen Befund ihrer früheren Untersuchung zum Erkennen der Reihenfolge dreier Töne auf (Divenyi & Hirsh, 1974), in der sie eine auffällig hohe Verwechslungshäufigkeit zwischen solchen Permutationen einer Tonfolge beobachteten, deren letzter Ton identisch war (z.B. zwischen tief-hoch-mittel und hoch-tief-mittel). Die Autoren vermuteten in

diesem Fall, daß die Identifikation der Reihenfolge der beiden ersten Töne durch die Anwesenheit des dritten Tons durch eine Art Auslöschung ("blanking") der Tonhöheninformation zumindest des zweiten Tons im echoischen Speicher zustandekommt.

Divenyi & Hirsh (1975) boten ihren Vpn drei 2o msec dauernde Töne (ISI = 0 msec) unterschiedlicher Frequenz dar, denen ein vierter, nicht zu beachtender Ton variabler Frequenz und Dauer ebenfalls mit einem ISI von 0 msec folgte. Die Aufgabe der Vpn bestand in der Identifikation der Reihenfolge der ersten drei Töne. Als Kontrollbedingung wurden die ersten drei Töne ohne nachfolgenden vierten Ton dargeboten.

Die Ergebnisse zeigen eine deutliche Überlegenheit der Identifikationsleistung (d') in der Kontrollbedingung gegenüber allen experimentellen Bedingungen. Die Stärke des "blanking effect" variierte in Abhängigkeit von 1. der Frequenz des vierten Tons relativ zu der der übrigen Töne, 2. von der Dauer des vierten Tons und 3. von der durch die vier Töne gebildeten Intervallkonfiguration. Die Beeinträchtigung der Identifikationsleistung gegenüber der Kontrollbedingung war dann am größten, wenn die Frequenz des letzten Tons mit der des ersten identisch war oder aber weit über der des höchsten der drei Töne lag. Die Identifikationsleistung wurde durch die Tondauer des letzten Tons dann am stärksten beeinträchtigt, wenn sie der aller vorausgehenden Töne entsprach. Wenn die Intervallkonfiguration in einer gleichgerichteten auf- oder absteigenden Intervallfolge bestand, war die Identifikationsleistung deutlich höher als bei allen anderen Konfigurationen. Die Autoren verglichen dieses Befundmuster mit den bekannten Maskierungsbefunden und konnten keinerlei Ähnlichkeiten feststellen.

Da die SOAs zwischen den Tönen in der Untersuchung von Divenyi & Hirsh (1975) in dem Bereich liegen, in dem im Prinzip starke rückwirkende Maskierungseffekte zu beobachten sein sollten (Massaro, 197o, 1972; Raab, 1963), mag das Ausbleiben typischer Maskierwirkungen ein wenig überraschen, weshalb die Autoren auch

annahmen, daß es sich hierbei um einen Interferenzeffekt auf dem Niveau des auditiven Kurzzeitgedächtnisses handeln müsse.

Wir stimmen in diesem Punkt nicht mit der Interpretation der Autoren überein, weil wir vermuten, daß es sich beim "blanking effect" um eine Variante des "suffix effect" (Crowder, 1978; Crowder & Morton, 1969) handelt und dieser auf Probleme des kognitiven Systems bei der Selektion des Suffix zurückzuführen ist, die bei der Informationsentnahme aus dem echoischen Speicher auftreten, wenn das Suffix oder der nicht zu beachtende vierte Ton bei Divenyi & Hirsh (1975) durch auditives Chunking zu einem Merkmalsverband verschmelzen.

Beim Suffix-Effekt handelt es sich um das erstmals von Conrad (1960) beobachtete Phänomen einer Beeinträchtigung der Reproduktion gesprochener Ziffernfolgen durch ein irrelevantes und nicht zu reproduzierendes letztes Item, das lediglich als Signal für den Beginn der Reproduktion dienen sollte. Crowder & Morton (1969) konnten zeigen, daß dieser Effekt verstärkt bei akustischer Darbietung auftritt, und Crowder & Raeburn (1970) demonstrierten die Unabhängigkeit des Effekts vom semantischen Gehalt des Suffix. Der die Reproduktion beeinträchtigende Effekt des Suffix tritt dann deutlich vermindert in Erscheinung, wenn das Suffix mit einer anderen Stimmqualität als die Items oder an anderer Stelle im Raum dargeboten wird (Crowder, 1973; Morton, Crowder & Prussin, 1971). Einen Suffix-Effekt für Tonfolgen hat Foreit (1976) demonstrieren können. Darwin & Baddeley (1974) wiesen nach, daß die Stärke des Effekts davon abhängt, wie gut die einzelnen Items voneinander diskriminierbar sind; bei guter Diskriminierbarkeit tritt ein geringerer Suffix-Effekt auf als bei schlechter.

Stimmqualität und insbesondere Lokalisation im Raum sind ausgezeichnete Selektionskriterien und Organisationsvariblen für die

auditive Wahrnehmung (Axelrod & Guzy, 1968; Guzy & Axelrod, 1972; ten Hoopen & Vos, 1979; vgl. auch Neumann, 1980). Auch Divenyi & Hirsh (1975) beobachteten einen geringeren Einfluß des vierten Tons auf die Identifikation der Reihenfolge der vorausgehenden Töne, wenn sich dieser deutlich durch seine längere Tondauer und/oder durch die Tonhöhe (allerdings nur im mittleren Bereich der vorgenommenen Frequenzvariationen) von den anderen Tönen abhob. Diese Befunde legen die Hypothese nahe, daß es sich beim Suffix- und "blanking"-Effekt um einen Aufmerksamkeitseffekt handeln könnte (Hitch, 1975; Kahneman, 1973), der darin besteht, daß das Suffix nicht von der Entnahme aus dem echoischen Speicher ausgeschlossen werden kann, wenn keine Selektionskriterien in Form von z.B. unterschiedlicher Stimmqualität, Lokalisation im Raum, relativer Dauer des Suffix gegenüber der der anderen Items oder durch einen Frequenzunterschied zu den anderen Tönen einer Tonfolge gegeben sind.

Die als Selektionskriterium dienenden physischen Dimensionen sind aber auch gleichzeitig diejenigen, die bei Ähnlichkeit der Merkmale einer Itemliste auf diesen Dimensionen ein auditives Chunking begünstigen. Der "blanking"-Effekt kann also dadurch erklärt werden, daß bei Ähnlichkeit der Merkmale des vierten Tons mit denen der vorausgehenden alle vier Töne ein Chunk bilden, aus dem der letzte Ton nicht ohne Fokussierung herausgetrennt werden kann. Wenn die Analyse aber auf Chunkniveau verbleibt und nicht die darunterliegende Elementebene sequentiell fokussiert, können aufgrund der globalen Chunkinformation leicht Fehler gemacht werden, weil der vierte Ton dann die Gesamtkonfiguration mitbestimmt. Wenn die Konfiguration der ersten drei Töne durch den vierten nicht tangiert wird, wie dies bei gleichgerichteten Intervallen zwischen allen vier Tönen der Fall ist, braucht der vierte Ton

nicht selegiert zu werden, weil die aufgabenadäquate Reaktion des Hörers sowohl bei Anwesenheit als auch Abwesenheit des vierten Tons in diesem Fall gleich ausfallen kann, so daß die aufgabenrelevante Information bereits auf Chunkniveau entnommen werden kann und sich damit eine fehleranfällige sequentielle Fokussierung der einzelnen Töne erübrigt.

Auch die Befunde von Watson et al. (1975) sind durch den Einfluß des auditiven Chunkings auf die Fokussierung einer zeitlich definierten Position und die Entnahme von Identitätsinformationen zu erklären. Die Autoren fanden in einer Wiedererkennungsaufgabe mit Tonfolgen, daß die Veränderung eines einzelnen Tons, dem eine kurze Pause folgt (während zwischen allen anderen Tönen keine Pausen liegen), mit größerer Wahrscheinlichkeit zu entdecken ist, als ohne nachfolgende Pause. Sie führten diesen Befund auf ein Nachlassen rückwirkender Maskierung zurück. Durch Einfügen einer Pause unmittelbar nach dem kritischen Ton wird aber die Tonfolge in zwei Teile segmentiert und als zwei Chunks repräsentiert. Statt die Verarbeitungseinheiten der gesamten Tonfolge auf Elementniveau zu analysieren, kann bei Einfügen einer Pause gleich das erste Chunk als Ganzes fokussiert werden, indem an der nun besonders ausgezeichneten Endposition bei Fokussierung des letzten Chunkelements der abweichende Ton besonders leicht aufgefunden und analysiert werden kann (vgl. ähnliche Befunde bei Divenyi & Hirsh, 1974; Leshowitz & Hanzi, 1974).

Da sich die Konzeption des auditiven Chunking offenbar bei der Erklärung scheinbar ganz unterschiedlicher Effekte zu bewähren scheint, wollen wir nun versuchen, das Problem der differentiellen

Schwierigkeiten bei der Identifikation der Reihenfolge von Tönen in kurzen Tonfolgen in Abhängigkeit von der Komplexität der Intervallkonfiguration auf dem Hintergrund des Hierarchiemodells und der Konzeption des auditiven Chunkings theoretisch zu lösen.

Zu erklären ist der Befund von Divenyi & Hirsh (1973, 1974, 1975), Idson & Massaro (1976), Kerivan & Carey (1976) und Ortmann (1926), wonach die Identifikation der Struktur von Drei- und Viertonfolgen mit Tönen verschiedener Frequenz dann am leichtesten ist, wenn die Tonfolge eine kontinuierlich auf- oder absteigende Konfiguration besitzt, und bei unregelmäßigem Intervallverlauf bedeutend schwieriger ist.

Dieser Befund ist im Rahmen des Hierarchiemodells dadurch zu erklären, daß die aufgabenrelevante Information bei einer kontinuierlich auf- oder absteigenden Tonfolge bereits auf dem Niveau der Gesamtrepräsentation (also dem höchsten Repräsentationsniveau) zu entnehmen ist, während dies für alle anderen Permutationen nicht gilt. Wenn wir annehmen, daß auf dem obersten Niveau lediglich die Konfiguration aller Töne ohne weitere Analyseschritte identifiziert werden kann, dann ist für eine Dreitonfolge, wie bei Divenyi & Hirsh (1973, 1974) und Idson & Massaro (1976), auf diesem Niveau entweder die Verlaufsform "kontinuierlich auf- oder absteigend" oder die Verlaufsform "Intervallverlaufsumkehrung" (von auf- nach absteigend oder von ab- nach aufsteigend) repräsentiert. Die Verlaufsform "kontinuierlich auf- oder absteigend" entspricht genau einer Reaktionskategorie, spezifiziert das Muster der Tonfolge also bereits eindeutig.

Aber die Verlaufsform "Intervallverlaufsumkehrung", von der es für jede Umkehrungsrichtung (auf-ab oder ab-auf) zwei Fälle gibt, nämlich großes Intervall zuerst und dann kleines oder umgekehrt, ist aufgrund der Information der Gesamtrepräsentanz noch nicht eindeutig auf eine Reaktionskategorie abbildbar. Deshalb muß bei dieser Verlaufsform das darunterliegende Repräsentationsniveau konsekutiver Intervalle fokussiert und die Intervalle sequentiell analysiert werden. Wenn man zusätzlich annimmt, daß jeder weitere Analyseschritt die Wahrscheinlichkeit einer fehlerhaften Identifikation erhöht, dann sind die Befunde zu den differentiellen Schwierigkeiten der Identifikation permutierter Dreitonfolgen erklärt.

Nimmt die Anzahl der Töne zu, dann nimmt auch, je nach Konfiguration der Intervalle, die Anzahl der Repräsentationsniveaus zu, die zwischen der Gesamtrepräsentation und der Elementrepräsentation liegen. Betrachten wir einmal das Beispiel einer Tonfolge aus sechs verschiedenen Tönen (1/2/3/4/5/6 sei die kontinuierlich aufsteigende Folge konstanter Intervalle). Die Identität der kontinuierlich auf- oder absteigenden Tonfolge kann (wie auch bei der Dreitonfolge) bereits aufgrund der Gesamtrepräsentation bestimmt werden. Die Folge 6/5/4/1/2/3 dagegen kann nicht auf dem obersten Niveau identifiziert werden, weil es noch eine weitere Permutation gibt, die ebenso aus einem kontinuierlich absteigenden Teil besteht, dem ein kontinuierlich aufsteigender folgt(nämlich 3/2/1/4/5/6). Auf dem unmittelbar unterhalb des Niveaus der Gesamtrepräsentanz liegenden Niveau kann diese Tonfolge als zwei Chunks repräsentiert sein, da die Ähnlichkeit der Tonhöhen innerhalb jedes Teils größer als zwischen ihnen ist und außerdem ein großer

Intervallsprung ihre Grenze deutlich markiert. Diese Tonfolge kann also nur über das Chunkniveau durch sequentielle Analyse der Chunks identifiziert werden.

Während also für die kontinuierlich auf- oder absteigende Tonfolge die Annahme zweier Repräsentationsniveaus ausreicht, muß bei der zuletzt beschriebenen Tonfolge von drei Repräsentationsniveaus ausgegangen werden, wobei zwischen der Gesamtrepräsentanz und den Elementrepräsentanzen das Niveau der Chunkrepräsentanzen liegt. Betrachten wir nun die Permutation 4/6/5/1/3/2: Sie kann auch noch nicht auf dem Niveau der Chunkrepräsentanzen identifiziert werden, weil es für jedes Chunk zwei Permutationen gibt, die dieselbe Teilkonfiguration besitzen (nämlich für das erste Chunk die Permutation 5/6/4 und für das zweite die Permutation 2/3/1). Bei dieser Tonfolge muß also die Binnenstruktur der Chunks erst durch sequentielle Analyse konsekutiver Intervalle identifiziert werden.

An diesem Beispiel wird deutlich, daß die Erklärung der Strukturidentifikation in Tonfolgen durch das Hierarchiemodell im Verbund mit dem Konzept des auditiven Chunkings im Prinzip auf beliebig lange und komplexe Tonfolgen übertragbar sein sollte. Dies sollte auch dann noch der Fall sein, wenn andere physische Merkmale als das der Ähnlichkeit der Tonhöhen für die Chunkbildung bestimmend werden, wie dies bei echten Melodien der Fall ist. Verallgemeinern wir die am Beispiel von Tonfolgen mit drei bzw. sechs Tönen gewonnenen Erkenntnisse über die kognitiven Analyseprozesse bei der Identifikation der Struktur einer Tonfolge, dann zeigt sich, daß das allgemeine Prinzip dieses Identifikationsprozesses darin besteht, daß die Gesamtkonfiguration vom höchsten Repräsentations-

niveau angefangen solange auf jeweils niedrigeren Repräsentationsniveaus analysiert werden muß, bis für jeden Ausschnitt der Gesamtkonfiguration ein Niveau erreicht ist, auf dem eine Teilkonfiguration als kontinuierlich auf- oder absteigende Verlaufsform repräsentiert ist. Dieses Erklärungsprinzip ist auch auf die Daten von Kerivan & Carey (1976) bei Viertonfolgen erfolgreich anwendbar.

Die bei dieser Darstellung der Prozesse als funktional wirksam vorausgesetzte Identitätskategorie für Intervallrichtungsverläufe ist die kontinuierlich auf- oder absteigende "tonale Linie". Daß es sich hierbei um eine grundlegende Identitätskategorie für tonale Muster handeln könnte, meint offenbar Ortmann (1933), wenn er die Schwierigkeit der Identifikation einer Tonfolge mit der Anzahl der Intervallrichtungswechsel in Beziehung bringt.

Beim strukturellen Hören können aber wahrscheinlich vom Hörer auch bedeutend komplexere Identitätskategorien herangezogen werden, die nicht in "elementarere" Identitätskategorien analytisch zerlegt zu werden brauchen. In unseren beiden ersten Experimenten werden wir einige zentrale Aussagen des dargestellten Erklärungsansatzes überprüfen; in Teil II überprüfen wir die Verallgemeinerbarkeit des Erklärungsansatzes für die Wahrnehmung echter Melodien, wobei sowohl strukturell einfache als auch bedeutend komplexere Identitätskategorien aufgabenrelevant werden.

4.4 ZUSAMMENFASSUNG

Das Hierarchiemodell von Neumann (1980) und Wolff (1977), das zu-

nächst nur als Rahmenkonzeption zur Erklärung funktionaler und struktureller Aspekte der kognitiven Verarbeitung und Repräsentation komplexer Reizmuster zu betrachten ist, bedurfte einer Spezifikation, die den Besonderheiten der Verarbeitung sich in der Zeit entwickelnder Reizmuster gerecht wird. Theoretisch zu lösen war das Problem, daß nicht alle Informationen über ein sich während eines größeren Zeitraums entwickelnden Reizmusters in gleicher Weise simultan repräsentiert sein können und daß deshalb ein reines "top-down"-Modell keine adäquate Erklärung der auditiven Mustererkennung zu geben vermag. Eine simultane Verfügbarkeit von Reizmusterinformationen ist nur innerhalb des als "psychische Präsenzzeit" bezeichneten gleitenden Zeitausschnitts anzunehmen.

Wir lösten dieses Problem, indem wir annahmen, daß die Verarbeitung einer Tonfolge dann nicht auf dem höchsten strukturellen Niveau innerhalb der Repräsentationshierarchie beginnt, wenn sich die Tonfolge über einen Zeitausschnitt erstreckt, der den der psychischen Präsenzzeit übersteigt. Dasjenige Repräsentationsniveau, dessen einzeln fokussierbare Verarbeitungseinheiten simultan verfügbar und als erste bewußt repräsentiert sind, bezeichneten wir als "primäres Repräsentationsniveau". Hierbei sollte es sich um ein Niveau handeln, auf dem als kleinste Einheiten Takte repräsentiert werden, es kann sich hierbei aber auch um Taktgruppen handeln, sofern das Tempo so hoch ist, daß zwei oder drei vollständige Takte den Zeitraum der psychischen Präsenzzeit nicht überschreiten.

In dieser Variante des Hierarchiemodells ist die Reizmusteranalyse nicht mehr, wie noch bei Wolff (1977), ausschließlich als

"top-down"-Verarbeitung zu betrachten, sondern integriert "top-down"- und "bottom-up"-Anteile, indem der Aufbau von supraprimären Repräsentanzen als "bottom-up"-Prozeß und die Detailanalyse primärer Repräsentanzen auf subprimären Niveaus als "top-down"-Prozeß konzipiert ist.

Die funktionale Bedeutung supraprimärer Repräsentationsniveaus diskutierten wir am Beispiel der "retrospektiven Umdeutung" eines vorausgegangenen musikalischen Ablaufs auf dem Hintergrund gegenwärtig fokussierter primärer oder subprimärer Repräsentanzen und am Beispiel der Erwartungsbildung aufgrund einer konfigurativen Repräsentation vergangener musikalischer Abläufe auf supraprimären Repräsentationsniveaus und ihrer Extrapolation in die Zukunft.

Schließlich übertrugen wir die Konzeption des "visuellen Chunkings" (Wolff, 1977) auf die auditive Modalität zur Erklärung einer strukturellen Binnengliederung primärer Repräsentanzen in Verarbeitungseinheiten, die oberhalb des Niveaus der Elementarrepräsentanzen anzusiedeln sind, um die in Kap. 2 erörterten Organisationseffekte tonaler Muster in das Modell einbeziehen zu können. Wir diskutierten dann experimentelle Befunde zur Wahrnehmung von Tonfolgen, die als empirische Belege für eine funktionale Interpretation des auditiven Chunkings als ein der echoischen Speicherung vorgeordneter Verarbeitungsprozeß zu betrachten sind, und erklärten auf diesem theoretischen Hintergrund scheinbar zusammenhanglose experimentelle Befunde.

Schließlich konnten wir das angestrebte Ziel erreichen, die experimentellen Befunde zu den differentiellen Schwierigkeiten der Identifikation von Tonfolgen in Abhängigkeit von der Intervallkonfiguration mit dem Hierarchiemodell und dem Konzept des auditiven Chunkings zu erklären.

5. EXPERIMENT 1: FOKUSSIERUNG STRUKTURELLER DETAILS EINER TONFOLGE IN ABHÄNGIGKEIT VOM VORAUSGEGANGENEN TONALEN KONTEXT

5.1 EINLEITUNG

Unser erstes Experiment dient der Überprüfung von Hypothesen, die sich aus den beiden zentralen Aussagen unseres Modells des Strukturerkennens in Tonfolgen ableiten lassen[1]. Die erste zentrale Aussage bezieht sich auf die Reihenfolge der Fokussierung einzelner Niveaus der strukturellen Repräsentationshierarchie:

Tonfolgen, deren zeitliche Erstreckung die der psychischen Präsenzzeit nicht überschreiten, werden strukturell analysiert, indem zuerst für die Repräsentanz der Gesamttonfolge geprüft wird, ob sie die aufgabenrelevante Information enthält oder nicht. Ist dies der Fall, dann kann bereits in diesem Stadium der Verarbeitung die Reaktionsvorbereitung einsetzen. Steht die aufgabenrelevante Information jedoch auf diesem obersten Repräsentationsniveau nicht zur Verfügung, werden die Verarbeitungseinheiten des unmittelbar darunterliegenden Niveaus fokussiert und es wird versucht, ihnen die aufgabenrelevante Information zu entnehmen. Führt auch dies nicht zur Einleitung der Reaktion, dann wird solange die jeweils niedrigere Repräsentationsebene durch den von Wolff (1977) beschriebenen Einschachtelungsprozeß fokussiert und analysiert, bis die aufgabenrelevante Information entnommen werden kann. Vorausgesetzt wird, daß die aufgabenrelevante Infor-

1) Wir stellen hier die Schlußfolgerungen und Hypothesen dar, die wir zum Zeitpunkt der Planung dieses Experiments (Winter 1977/78) für zutreffend hielten. Es wird sich später im Verlauf der Diskussion der Befunde herausstellen, daß sie nicht in allen Punkten konsequent durchdacht waren.

mation diejenige Information über ein strukturelles Detail ist, die als "tonale Linie" identifiziert werden kann.

Die zweite zentrale Aussage bezieht sich auf das Konzept des auditiven Chunkings und seine Rolle im Rahmen des Hierarchiemodells:

In einem der echoischen Speicherung vorausgehenden Verarbeitungsschritt findet eine Gruppierung von Tönen nach den Prinzipien der Ähnlichkeit ihrer physischen Merkmale und zeitlichen Nähe statt, die zu Verarbeitungseinheiten (Chunks) führt, deren vertikale Position innerhalb der strukturellen Repräsentationshierarchie zwischen dem Niveau der Gesamtrepräsentanz der Tonfolge und dem Niveau seiner Elementarrepräsentanzen liegt.

Den methodologischen Weg zur Überprüfung von Hypothesen, die aus diesen Aussagen im Kontext einer gegebenen Strukturidentifikationsaufgabe abzuleiten sind, weist uns Garner (1978):

"There are strong suggestions that the emergent configurational properties tend to dominate perception of stimuli, overriding component properties, when they are available" (Garner, 1978, S. 124). "Are configurational properties, then, somehow psychologically more fundamental? ... If there is some reasonable way to decide which is primary and which is secondary, then that differentiation would be useful. My own inclination is to consider as primary those properties, that are processed under a speed stress" (Garner, 1978, S. 13o-131).

Die Verarbeitungsreihenfolge bei Durchlaufen der Repräsentationshierarchie können wir also nur dann rekonstruieren, wenn wir den Vpn eine Strukturidentifikationsaufgabe geben, die eine Reaktions-

zeitmessung sinnvoll erscheinen läßt. Das ist dann der Fall, wenn die zu erwartende Fehlerzahl gering und die Reaktionsalternativenzahl klein (möglichst zwei) ist und unabhängig von der Struktur der Tonfolge konstant bleibt, d.h. wir müssen strukturell äquivalente Tonfolgen verwenden, die einmal eine richtige Reaktion mit der einen und das andere Mal mit der anderen Reaktionsalternative erlauben.

Eine solche Aufgabe ist die folgende einfache Konturidentifikationsaufgabe: Es werden drei Töne dargeboten, von denen der erste Ton nicht zu beachten ist. Aufgabe der Vpn ist es, die Intervallrichtung der beiden letzten Töne als auf- oder absteigend zu klassifizieren (Zwei-Alternativen-Wahlreaktionsaufgabe). Wird z.B. eine Tonfolge verwendet, die aus einem ersten aufsteigenden Intervall von drei Halbtönen und einem zweiten absteigenden Intervall von einem Halbton besteht, dann ist die strukturäquivalente Tonfolge, die eine Reaktion mit der anderen Reaktionsalternative verlangt, die Tonfolge, die mit einem absteigenden Intervall von drei Halbtönen beginnt und mit einem aufsteigenden Intervall von einem Halbton endet. Der Konturverlauf oder die Konfiguration dieser Tonfolgen ist bis auf die Spiegelung an der Zeitachse gleich (vgl. Abb. 2).

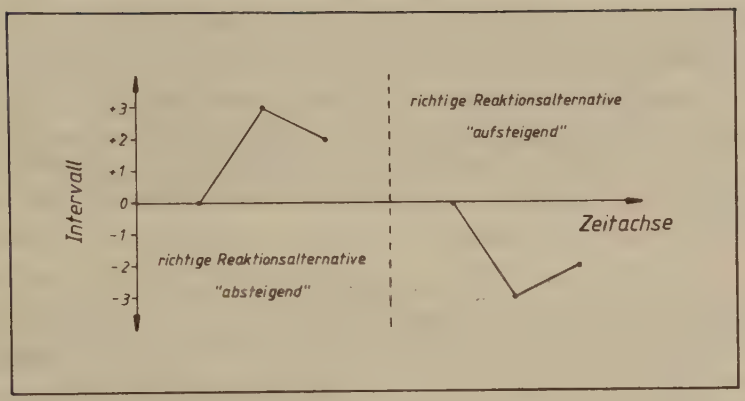


Abb. 2: Beispiel zweier strukturell äquivalenter Tonfolgen (Beschreibung siehe Text)

Obwohl die Vpn nur den Konturverlauf des zweiten Intervalls beachten sollen, muß sich, wenn unser Modell eine zutreffende Beschreibung des Analyseprozesses gibt, ein Effekt des vorausgehenden Tons nachweisen lassen, der darin bestehen sollte, daß der erste Ton als integraler Bestandteil der Gesamtrepräsentation der Tonfolge mit in die Entscheidung einfließt, ob die aufgabenrelevante Information bereits auf dem Niveau der Gesamtrepräsentanz oder erst auf einem niedrigeren Niveau verfügbar ist. Entsprechend unserer Argumentation bei der Erklärung der differentiellen Schwierigkeiten der Identifikation einer Tonfolge sollte

auch in dieser Aufgabensituation bei einer kontinuierlich auf- oder absteigenden Gesamtkonfiguration die Information über die Intervallrichtung der beiden letzten Töne bereits der Gesamtrepräsentanz zu entnehmen sein, während bei gegensinnigen Richtungen der beiden Intervalle diese Information erst auf einem niedrigeren Repräsentationsniveau zu entnehmen sein dürfte: Bei gleichsinnig verlaufenden Intervallen sollte die Reaktionszeit (RT) also kürzer als bei gegensinnigem Verlauf sein.

Variiert man außer der Intervallkonfiguration auch noch die Ähnlichkeit der physischen Merkmale der Tonkomponenten sowie das Verhältnis der Pausen zwischen ihnen, dann sollten diese Variationen Bedingungen darstellen, die zu unterschiedlichen Chunkorganisationen auf dem Repräsentationsniveau direkt unterhalb des der Gesamtrepräsentanz führen. Wenn das erste gegenüber dem zweiten Intervall relativ groß oder die Pause zwischen den ersten beiden Tönen relativ lang gegenüber der Pause zwischen dem zweiten und dritten Ton ist, sollten die beiden letzten Töne ein Chunk bilden, das bei gegensinnigem Konturverlauf, also bei Wechsel der Repräsentationsebene, direkt fokussiert werden kann, während im umgekehrten Fall, nämlich bei relativ kleinem ersten Intervall im Verhältnis zum zweiten oder bei relativ kurzer erster Pause gegenüber der zweiten, die beiden ersten Töne eine Verbindung zu einem Chunk eingehen, so daß die aufgabenrelevante Information aus den Verarbeitungseinheiten auf Chunkniveau nicht entnommen werden kann, sondern erst auf dem Niveau der Elementrepräsentanzen. Wenn also eine Chunkorganisation vorliegt, die die beiden letzten Töne zu einer Verarbeitungsein-

heit vereinigt, sollte die RT kürzer sein, als wenn die ersten beiden Töne eine Verarbeitungseinheit bilden.

Wenn kein Chunkniveau vorliegt bzw. dieses mit dem der Gesamtrepräsentanz zusammenfällt, dann sollte die RT zwischen den Werten liegen, die bei den gerade beschriebenen Chunkorganisationen zu erwarten sind: In diesem Fall muß nämlich die Information über die Kontur des letzten Intervalls durch sukzessive Analyse der drei Elementrepräsentanzen gewonnen werden, wie dies auch dann der Fall ist, wenn die beiden ersten Töne ein Chunk bilden und deshalb auf das Niveau der Elementarrepräsentanzen übergegangen werden muß. Aber im Falle der Identität von Gesamtrepräsentanz und Chunkrepräsentanz kann die Elementebene direkt fokussiert werden, während im Falle der die Entnahme der aufgabenrelevanten Information behindernden Chunkorganisation die Elementebene nur über die der Chunks fokussiert werden kann, also ein zusätzlicher Verarbeitungsschritt benötigt wird.

Die gerade beschriebenen Effekte der Chunkorganisation sollten sich aber deutlich nur für Tonfolgen mit gegensinnigem Konturverlauf zeigen, da nur in diesem Fall die Notwendigkeit besteht, die aufgabenrelevante Information auf einem Niveau unterhalb des der Gesamtrepräsentanz zu entnehmen.

5.2 METHODE

Reizmaterial

Die Tonfolgen bestanden aus drei reinen Sinustönen von 6o msec Dauer mit weichem Amplitudenanstieg und -abfall (Dämpfungsabfall von ca. 6o db (A) in 2o msec, Dämpfungsanstieg um ca. 6o db (A) in 4o msec) und einer Spitzenamplitude von 72 db (A) (gemessen am Kopfhörer).

Die Töne wurden aus dem drei Oktaven umfassenden Frequenzbereich zwischen 22o Hz (a, eine Oktave tiefer als der Kammerton a') und 172o Hz genommen und gehörrichtig lautstärkenkorrigiert. Die Frequenz des jeweils ersten Tons wurde mit Hilfe eines gleichverteilte Zufallszahlen produzierenden Zufallsgenerators unter der Einschränkung gewählt, daß der tiefste bzw. höchste Ton der für eine experimentelle Bedingung zu konstruierenden Tonfolge die abgesteckten Frequenzgrenzen nicht überschreiten darf. Auf diese Weise erhielten wir für jeden Ton eine Gleichverteilung der Frequenzen über alle Versuchsbedingungen hinweg, so daß eine besonders auffällige hohe oder tiefe Frequenz des letzten Tons kein Kriterium für eine Reaktionsentscheidung sein konnte.

Versuchsapparaturen

Die Töne wurden von einem digital steuerbaren Tongenerator (Präzisions-NF-Generator von Rohde & Schwarz) produziert, der on-line durch einen Prozeßrechner (Dietz 621/8) gesteuert wurde

(Genauigkeit der zeitlichen Steuerung: 1 msec). Die RT-Messung wurde ebenfalls vom Prozeßrechner mit 1 msec Genauigkeit durchgeführt, wenn eine der beiden Morsetasten einen Interrupt-Eingang aktivierte, mit denen die Vpn ihre Reaktionen abgaben.

Die Amplitudendämpfung wurde durch einen knackfreien Schalter vorgenommen (zwei gegengekoppelte photoelektrische Widerstände mit schnellen Glühlämpchen, Eigenbau des Psychologischen Instituts der Ruhr-Universität Bochum), der zwischen Vorverstärkerausgang am Tongenerator und Kopfhörer (Sennheiser HD 42o) geschaltet war.

Versuchsplan

Der Versuchsplan umfaßt fünf experimentelle Variablen:

1. Intervallkonfiguration: gleichsinniger vs. gegensinniger Intervallverlauf (alle Intervalle auf- oder absteigend vs. ein Intervall aufsteigend (absteigend) und das andere absteigend (aufsteigend));
2. Intervallgröße zwischen 1. und 2. Ton ($\Delta f_{1,2}$): 0, 1/2, 1, 2, 4, 8 oder 16 Halbtöne;
3. Intervallgröße zwischen 2. und 3. Ton ($\Delta f_{2,3}$): 1, 2, 4 oder 8 Halbtöne;
4. SOA zwischen 1. und 2. Ton ($SOA_{1,2}$): 9o, 18o, 36o oder 72o msec sowie die Kontrollbedingung, in der nur zwei Töne dargeboten werden, die als $SOA_{1,2} \rightarrow \infty$ betrachtet wird;
5. SOA zwischen 2. und 3. Ton ($SOA_{2,3}$): 18o oder 36o msec.

Die Variable "Richtung des kritischen Intervalls" wurde als Meßwiederholung betrachtet. Aus der unabhängigen Variation aller experimentellen Variablen resultiert ein Versuchsplan mit 2 (Konfigurationen) X 7 ($\Delta f_{1,2}$) X 4 ($\Delta f_{2,3}$) X 5 ($SOA_{1,2}$ + Kontrollbedingung) X 2 ($SOA_{2,3}$) = 56o Bedingungen, die viermal dargeboten und dabei in jeweils 5o % der Fälle mit aufsteigendem oder absteigendem kritischen Intervall kombiniert wurden.

Versuchsdurchführung

Für jede Vp wurde mit Hilfe des Prozeßrechners durch zufallszahlenabhängiges Mischen eine individuelle Reihenfolge der Versuchsbedingungen erstellt, die aber nicht die Kontrollbedingung einschloß. Diese wurde, von Vp zu Vp zufällig gewählt, als Block entweder zu Beginn, in der Mitte oder am Ende einer Versuchssitzung durchgeführt.

Jede der beiden Versuchssitzungen begann mit einem Training im Erkennen der Richtung des letzten Intervalls. In der ersten Sitzung wurden 4oo Trainingstonfolgen, in der zweiten 1oo dargeboten. Hierbei hörte der Versuchsleiter über einen zweiten Kopfhörer mit und gab der Vp Rückmeldung, zuerst bei jeder einzelnen Tonfolge, später nur noch im Fehlerfall. Eine Versuchssitzung dauerte ca. 3 Stunden und wurde (nach jeweils 4o Minuten) von zwei halbstündigen Pausen unterbrochen.

Ein einzelner Versuchsdurchgang wurde durch einen 2oo msec dauernden Warnton von 25oo Hz eingeleitet, dem eine Pause von 1.5 sec

folgte, bevor die Tonfolge dargeboten wurde. Die RT-Messung wurde vom Prozeßrechner unmittelbar nach Ausschalten des dritten Tons gestartet und durch Tastendruck (rechts für "aufsteigend", links für "absteigend" in der ersten Versuchssitzung, in der zweiten umgekehrt) der Vp gestoppt. Nach Abgabe der Reaktion trat eine Pause von 3.5 sec ein, bevor der nächste Durchgang durch einen Warnton angekündigt wurde. Die Richtigkeit der Reaktion überprüfte der Prozeßrechner, und im Fehlerfall wurde dieser Versuchsdurchgang sowie der unmittelbar folgende an einer für die Vp nicht vorhersehbaren Stelle wiederholt. Nach Abschluß des Trainings gaben wir keine Rückmeldungen mehr.

Instruktion

Die Instruktion beschrieb der Vp die Aufgabe und betonte, daß sie den ersten Ton einer Tonfolge nicht zu beachten brauche. Hervorgehoben wurde, daß sie so schnell wie möglich die richtige Taste drücken, aber Fehler unbedingt vermeiden solle. Eine schnelle Reaktionen ermöglichende Fingerhaltung über den Tasten wurde ihr demonstriert, sofern sie diese nicht schon spontan einnahm. In einigen wenigen Fällen mußte die schriftlich vorgelegte Instruktion im Verlauf des Trainings der ersten Versuchssitzung durch mündliche Erklärungen ergänzt werden, wenn die Vpn nicht zwischen den Konturverläufen differenzieren konnten. In diesen Fällen wurden einige dargebotene Tonfolgen graphisch veranschaulicht, was dann sehr schnell zu beinahe völliger Fehlerfreiheit führte.

Versuchspersonen

Als Versuchspersonen nahmen 14 musikalisch trainierte Vpn (mit mehr als drei Jahren Unterricht im Spielen eines Musikinstruments) und 16 musikalisch untrainierte Vpn teil. Bei den Vpn handelte es sich um Studenten der Fachrichtung Psychologie, die im Rahmen ihrer Versuchspersonenverpflichtungen am Experiment teilnahmen.

5.3 BEFUNDE UND DISKUSSION

Auswertungsverfahren

Wir rechneten mit der abhängigen Variablen "Reaktionszeit" eine Varianzanalyse[1)] für abhängige Messungen auf allen fünf experimentellen Bedingungsvariablen, einmal ohne Einbeziehung der Kontrollbedingung (s. Anh. 3/1.1) und einmal mit der Kontrollbedingung als fünfte Stufe des Faktors $SOA_{1,2}$ (s. Anh. 3/1.2). Wegen der hohen Wahrscheinlichkeit inhomogener Kovarianzen zwischen Bedingungen in einem Design mit abhängigen Messungen (Box, 1953) verwandten wir zur Schätzung der Irrtumswahrscheinlichkeit der F-Werte ein kombiniertes Kriterium: Als signifikant wird ein F-Wert bezeichnet, der bei gegebenen Freiheitsgraden eine Irrtumswahrscheinlichkeit kleiner oder gleich 0.02 besitzt und für den gleichzeitig die konservative Schätzung nach Greenhouse & Geisser (1959) (bei df = 1; N - 1) kleiner oder

---

1) Programm ANOVA 5 des Autors; die Fehlervarianzen wurden wegen zu erwartender Inhomogenität für jeden Effekt einzeln berechnet (also kein gepoolter Fehler).

gleich 0.10 beträgt; als sehr signifikant bezeichnen wir einen F-Wert, dessen Irrtumswahrscheinlichkeit nach beiden Kriterien kleiner oder gleich 0.01 ist. Mittelwertdifferenzen wurden mit Scheffé-Kontrasten überprüft (vgl. Morrison, 1967).

Eine Fehlerrate wurde zwar auch erhoben, da sie aber mit 2.4 % recht niedrig liegt, erscheint es wenig sinnvoll, eine separate Fehleranalyse vorzunehmen, denn die durchschnittlich 54 Fehler verteilen sich auf die 56o Bedingungen einigermaßen gleichmäßig, so daß die Fehlersumme pro Bedingung über alle Vpn hinweg gerade eine Häufigkeit von 3 erreicht.

Vergleich der experimentellen und Kontrollbedingungen

Die Abhängigkeit der Identifikation einer Intervallrichtung vom vorausgehenden Kontext, der, wenn dies für das kognitive System möglich wäre, unbeachtet bleiben könnte, demonstriert eine Gegenüberstellung der Kontrollbedingungen und experimentellen Bedingungen. Wie Abb. 3 zeigt, liegt der Mittelwert aller Kontrollbedingungen (über die $\Delta f_{2,3}$- und $SOA_{2,3}$ - Bedingungen gemittelt) mit einer RT von 395 msec deutlich unter dem aller experimentellen Bedingungen mit einer RT von 6o2 msec (Scheffé-Test für die Differenz zwischen dem Mittelwert aller $SOA_{1,2}$ - Bedingungen und den Kontrollbedingungen: $p<0.01$; $SOA_{1,2}$ - Haupteffekt: $F(4,116)=5o.47$, $p<0.01$; bei df=1,29 ebenfalls $p<0.01$).

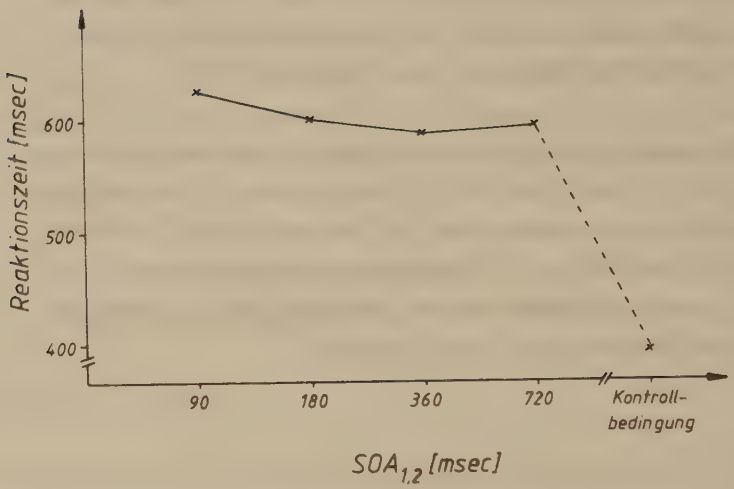


Abb. 3: RT-Mittelwerte der $SOA_{1,2}$ - Bedingungen und der Kontrollbedingung

Dem Hörer ist es offenbar nicht möglich, den ersten Ton in den experimentellen Bedingungen von einer kognitiven Analyse auszuschließen, und dies auch dann nicht, wenn der zeitliche Abstand zwischen dem ersten und den beiden folgenden Tönen relativ groß ist, wie dies bei $SOA_{1,2}$ = 72o msec der Fall ist, obwohl in diesem Fall ein ausgezeichnetes Selektionskriterium vorhanden sein könnte. Ein Modell, das davon ausgeht, daß einzelne Töne dem echoischen Speicher entnommen werden, wie das von Massaro (1975, 1976a), kann diesen Befund nicht erklären. Auch das mögliche Argument, die Verarbeitung von drei Tönen dauere länger als die von nur zwei Tönen, ist nicht stichhaltig, denn nach Massaro (197o)

dauert die Verarbeitung eines einzelnen Tons nur 25o msec, so daß bei einem $SOA_{1,2}$ von 72o msec (ISI = 66o msec) die Verarbeitung des ersten Tons bereits lange abgeschlossen sein müßte, wenn der zweite Ton erklingt. Der Unterschied zwischen den experimentellen und den Kontrollbedingungen kann also nur dadurch erklärt werden, daß a u c h d e r e r s t e T o n genau wie die anderen verarbeitet worden ist, b e v o r eine Selektion wirksam werden konnte. Dies erfordert aber die Annahme, daß nicht nur einzelne Töne, sondern V e r b ä n d e v o n T ö n e n dem echoischen Speicher entnommen werden können. Die Gesamttonfolge muß also zeitlich vor den sie konstituierenden Einzeltönen repräsentiert worden sein.

Der Einfluß der Gesamtkonfiguration auf die Fokussierbarkeit des letzten Intervalls

Wenn die Gesamttonfolge zuerst repräsentiert ist, dann sollte diese Repräsentation auch die Informationsgrundlage für die Fokussierung des letzten Intervalls darstellen, d.h. die Schwierigkeit der Fokussierung sollte davon abhängen, in welche Konfiguration das Intervall eingebettet ist.

Nun ist es aber denkbar, daß in einer Situation mit Zeitdruck immer dann eine bezüglich des Verarbeitungsaufwands ökonomischere Strategie verfolgt wird, wenn dies möglich ist. Eine solche ökonomischere Strategie bestünde darin, die Stufe der Selektion zu überspringen und die aufgabenrelevante Information der zuerst zur Verfügung stehenden Repräsentation direkt zu entnehmen. Die Rich-

tung des letzten Intervalls kann aber nur dann der globalen Repräsentanz des Gesamtmusters entnommen werden, wenn die Richtung des ersten mit der des zweiten Intervalls übereinstimmt, so daß die auf dem höchsten strukturellen Niveau repräsentierte Konfiguration einen kontinuierlich auf- oder absteigenden Verlauf zeigen muß.

Der Haupteffekt der Intervallkonfiguration zeigt, daß unsere Überlegungen in diesem Punkt richtig sein müssen (s. Abb. 4). Der RT-Mittelwert für einen gleichsinnigen Konturverlauf ist mit 546 msec hochsignifikant niedriger als der für einen gegensinnigen Konturverlauf, bei dem ein Mittelwert von 659 msec erreicht wurde ($F(1,29) = 24.91$, $p \leq 0.01$).

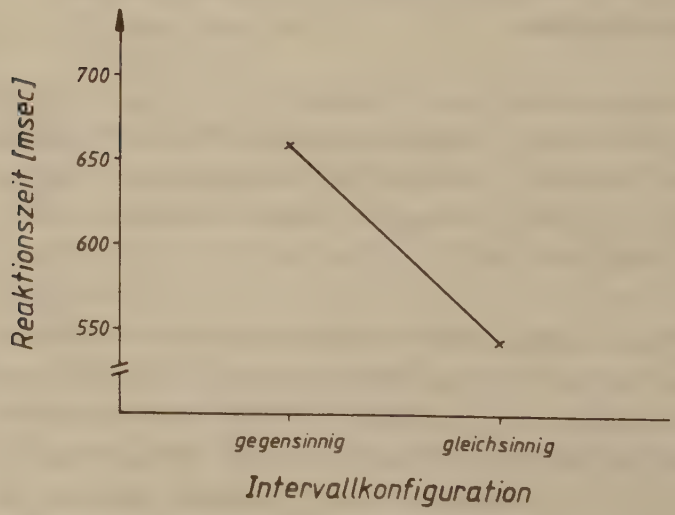


Abb. 4: RT-Mittelwerte für gegensinnige und gleichsinnige Intervallkonfiguration

Das Bild wird aber bedeutend differenzierter, wenn wir sozusagen auf das "nächst niedrigere Repräsentationsniveau der Daten" übergehen und die Interaktionen zwischen dem Faktor "Intervallkonfiguration" und den einzelnen Intervallen $\Delta f_{1,2}$ und $\Delta f_{2,3}$ betrachten.

Abb. 5 zeigt, daß die RT mit größer werdendem $\Delta f_{1,2}$ zunimmt und daß diese Zunahme bei gegensinniger Intervallkonfiguration bereits bei relativ kleinem $\Delta f_{1,2}$ beginnt, während sie bei gleichsinniger Intervallkonfiguration erst bei deutlich größeren Intervallwerten für $\Delta f_{1,2}$ zu steigen beginnt ($F(6,174) = 26.53$, $p < 0.01$; bei df = 1,29 ebenfalls $p < 0,01$).

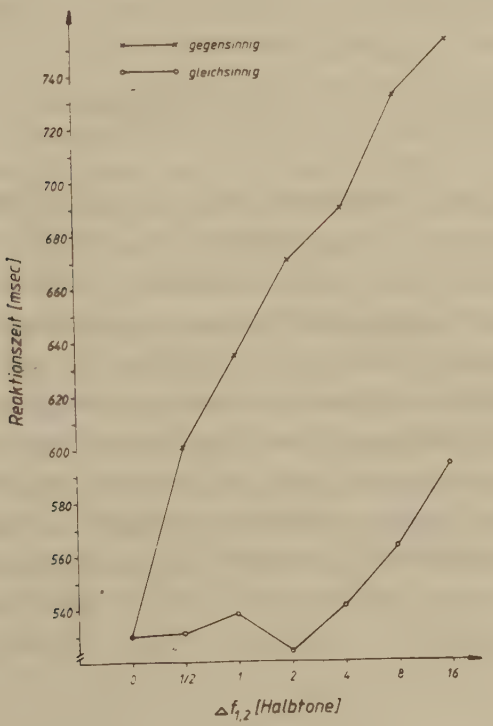


Abb. 5: RT-Mittelwerte der Interaktion "Intervallkonfigu-

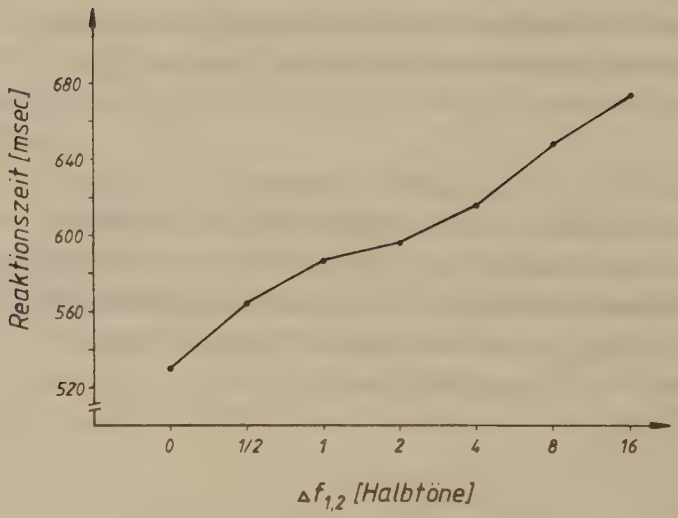


Abb. 6: RT-Mittelwerte für den $\Delta f_{1,2}$ - Haupteffekt

Die Zunahme der RT in Abhängigkeit von $\Delta f_{1,2,}$ die sich auch als Haupteffekt zeigt ($F(6,174) = 35.98$; $p<0.01$; bei df = 1,29 ebenfalls $p<0.01$; s. Abb. 6), stellt ein zunächst überraschendes Ergebnis dar, denn unsere Hypothese drückte eine Erwartung in genau der umgekehrten Richtung aus. Wir nahmen an, daß alle physischen Merkmale der Elementrepräsentanzen Bedingungen darstellen sollten, die ein auditives Chunking bewirken, das bei geringer Ähnlichkeit der Merkmale des ersten und zweiten Tons, aber höherer Ähnlichkeit der Merkmale des zweiten und dritten Tons eine Binnengliederung der Gesamtrepräsentanz entstehen lassen, wodurch zwischen Gesamtrepräsentanz und Elementarrepräsentanzen ein weiteres Niveau von Verarbeitungseinheiten konstitu-

iert wird, das das kritische letzte Intervall als direkt fokussierbare Verarbeitungseinheit ausgliedert. Diese Ausgliederung sollte gerade bei gegensinnigem Konturverlauf förderlich sein, da in diesem Fall ja die aufgabenrelevante Information nicht auf dem Niveau der Gesamtrepräsentanz entnommen werden kann.

Nun aber zeigt sich, daß ein großes $\Delta f_{1,2}$ eine die RT verlängernde Wirkung besitzt, die außerdem auch noch bei gleichsinnigem Verlauf in Erscheinung tritt. In Grenzen ist der RT-Anstieg bei Zunahme von $\Delta f_{1,2}$ und gegensinnigem Konturverlauf aber als selbstverständlich zu betrachten, denn das konfigurative Merkmal "Gegensinnigkeit" tritt ja bei $\Delta f_{1,2} > 0$ überhaupt erst auf. Was aber nicht erwartet wurde, ist eine Zunahme der RT über einen mittleren Wert von $\Delta f_{1,2}$ hinaus. Vielmehr erwarteten wir, daß zwar bei kleinen bis mittleren Werten von $\Delta f_{1,2}$ für gegensinnigen Konturverlauf eine RT-Erhöhung eintritt, die aber dann wieder bei weiterem Anstieg von $\Delta f_{1,2}$ in eine Abnahme umschlagen sollte, wenn $\Delta f_{1,2}$ so groß geworden ist, daß sich durch auditives Chunking der zweite und dritte Ton vom ersten deutlich abhebt. Ebenfalls nicht erwartet wurde der RT-Anstieg bei gleichsinnigem Konturverlauf, wie er sich in Abb. 5 von $\Delta f_{1,2} = 4$ Halbtönen an zeigt. Bei gleichsinnigem Konturverlauf sollte sich $\Delta f_{1,2}$ nach unserer Erwartung überhaupt nicht auf die RT auswirken, da die bei großem $\Delta f_{1,2}$ eventuell entstehende Binnengliederung der Gesamtrepräsentation eine Entnahme der aufgabenrelevanten Information auf diesem Niveau nicht behindern sollte.

Der $\Delta f_{1,2}$ - Effekt dürfte in dieser Form dann nicht auftreten, wenn er funktional mit dem auditiven Chunking in Beziehung zu bringen wäre. Wenn also nicht das Chunkniveau betroffen ist, kann sich der $\Delta f_{1,2}$ - Effekt eigentlich nur auf das Niveau der Gesamtrepräsentation ausgewirkt haben. Diese Auswirkung müßte aber dann darin bestehen, daß bei großem $\Delta f_{1,2}$ eine Fokussierung der beiden letzten Töne gegenüber kleinem $\Delta f_{1,2}$ verzögert wird.

Eine Verzögerung der Fokussierung des letzten Intervalls wäre z.B. dann denkbar, wenn die Gesamtrepräsentanz n i c h t zeitlich vor einer strukturell niedrigeren Repräsentation der Tonfolge vorliegt. Wenn diese Argumentation zutreffen sollte, dann müßte die Gesamtrepräsentanz erst durch Integration elementarerer Repräsentanzen aufgebaut werden, um ihre Steuerungsfunktion für den Fokussierungsvorgang auf einem strukturell niedrigeren Niveau vornehmen zu können. Ein solcher Integrationsprozeß könnte die funktionale Ursache für die Verzögerung der RT sein.

Warum ist aber die aufgabenrelevante Information nicht auch ohne diesen "Umweg" über die Integration einer Gesamtrepräsentanz dem echoischen Speicher zu entnehmen? Die Antwort hierauf ist auf dem Hintergrund des Hierarchiemodells eigentlich ganz einfach. Die aufgabenrelevante Information ist eine über die Teilkonfiguration der Tonfolge, die durch die Beziehung der beiden letzten Töne gegeben ist. Diese Beziehung muß aber zuerst in irgendeiner Form repräsentiert sein, bevor sie in eine ablaufende Handlung eingebracht werden kann. Wenn aber eine Tonfolge nicht

auf einem strukturell höheren Niveau als Ganzes repräsentiert ist, dann kann sie - wenn wir das hier offenbar nicht funktional werdende Chunkniveau außer Betracht lassen - nur noch in Form ihrer einzelnen Elemente repräsentiert sein, aber diese Elementrepräsentanzen enthalten nicht die aufgabenrelevante Information über die Teilkonfiguration der beiden letzten Töne. Es muß dann erst eine strukturell höhere Repräsentation erstellt werden, die eine Fokussierung dieser Teilkonfiguration erlaubt. Hierbei kann es sich nur um ein Repräsentationsniveau handeln, dessen strukturelles Niveau so hoch ist, daß es die zu fokussierende Teilkonfiguration als Verarbeitungseinheit enthält.

Die Teilkonfiguration der beiden letzten Töne kann nicht nach einer Integration n u r d i e s e r b e i d e n T ö n e entnommen werden, denn diese müßten ja dann auch sequentiell fokussierbar sein, was sie aber nicht sein können, wenn ihnen eine übergeordnete Repräsentanz fehlt, die sie sozusagen als Merkmal des Objekts "Tonfolge" trägt. Die Fokussierung eines Merkmals bedarf immer eines Objekts, an dem es als Merkmal überhaupt erst in Erscheinung tritt (vgl. auch die Aufmerksamkeitstheorie von Neumann, 1980).

Wenn es keine die drei Töne als Elemente oder Merkmale enthaltende Repräsentanz auf einem übergeordneten Niveau gäbe, wären die Töne auch nicht als zeitlich geordnet repräsentiert, denn das kognitive System kann Zeit nicht wahrnehmen, sondern nur zeitliche Abläufe aus einer B e z i e h u n g zwischen Ereignissen "erschließen". Zeit ist nicht als solche in abstrakter

Form kognitiv repräsentierbar, sondern nur in nicht-zeitlicher Form als Beziehungsstruktur. Die funktionale Notwendigkeit einer strukturell übergeordneten Repräsentation ergibt sich also daraus, daß auch die Fokussierung der Elementarrepräsentanzen eine zeitliche Beziehung erforderlich macht, die in nicht-zeitlicher Form kodiert ist.

Eine alternative Erklärung für den $\Delta f_{1,2}$ - Effekt, der wir eher zuneigen, als der ersten, besteht darin anzunehmen, daß zwar als erste Repräsentanz i m m e r eine Gesamtrepräsentanz gebildet wird, sich diese aber bei sehr großem $\Delta f_{1,2}$ bezüglich der verfügbaren Information von einer Gesamtrepräsentanz unterscheidet, wie sie bei kleinem $\Delta f_{1,2}$ gebildet wird. Der Streamingeffekt (Bregman, 1978a,b), der ja phänomenal darin besteht, daß die Beziehung zwischen sukzessiven Tönen nicht mehr wahrgenommen werden kann, wenn das Intervall zwischen ihnen relativ zur Länge des Zeitintervalls, in dem es auftritt, sehr groß ist, beruht nach Bregman (1978a) auf präattentiven Mechanismen der Gruppierung nach Ähnlichkeit. Umgekehrt kann man aber auch sagen, er bestehe in einer T r e n n u n g d e s U n ä h n l i c h e n . Wenn wir den $\Delta f_{1,2}$ - Effekt so betrachten, erscheint es plausibel anzunehmen, daß der qualitative Unterschied zwischen einer Gesamtrepräsentanz bei kleinem $\Delta f_{1,2}$ und großem $\Delta f_{1,2}$ darin besteht, daß in letzterem Fall keine konfigurativen Merkmale mehr repräsentiert sind, die sich auf alle drei Töne erstrecken. Die über die Gesamtrepräsentanz in diesem Fall noch verfügbare Information ist dann lediglich die über die zeitliche Sukzession eines isolierten Tones und eines Tonpaares. Die Fokussierung des letzten

Intervalls ist dann alleine aufgrund der zeitlichen Information möglich. Dieser Informationsverlust könnte sich nun dahingehend auswirken, daß die Steuerung der Fokussierung ungenauer wird und deshalb mehr Zeit in Anspruch nimmt.

Im Effekt von $\Delta f_{1,2}$ auf die RT müssen wir demnach das auditive Äquivalent eines V e r l u s t e s v o n O b j e k t - p e r m a n e n z sehen. Die relative Stabilität der Beziehungen zwischen den physischen Merkmalen einer Tonfolge stiftet ihre Permanenz als Wahrnehmungsgegenstand. Relative Stabilität liegt aber nur dann vor, wenn die Veränderung der physischen Merkmale bezogen auf den Zeitraum, in dem diese Veränderung stattfindet, d.h. die erste Ableitung einer Funktion der Merkmalsvariation nach einem Zeitintervall Δt, für ein bestimmtes Repräsentationsniveau einigermaßen konstant bleibt. Diese Formulierung ist nichts anderes als eine mathematische Abstraktion des I n v a r i a n z k o n z e p t s von Gibson (195o, 1966a,b).

Nun wird auch verständlich, warum der Faktor $\Delta f_{1,2}$ mit dem der Intervallkonfiguration interagieren muß (s.Abb. 5): Wenn man von der Variation der SOAs zwischen Tönen und der von $\Delta f_{2,3}$ abstrahiert, dann liegt bei gegensinnigem Konturverlauf eine um so größere Variabilität des Merkmals "Tonhöhe" vor, je größer $\Delta f_{1,2}$ ist, die bei jedem Wert von $\Delta f_{1,2}$ bei gegensinnigem Verlauf größer als bei gleichsinnigem ausfällt. Da die Objektpermanenz aber nicht schon bei geringfügiger Merkmalsvariabilität verloren geht, muß $\Delta f_{1,2}$ für gleichsinnigen Konturverlauf erst einen bestimmten Schwellenwert übersteigen, bevor auch bei dieser Intervallkonfiguration die Permanenz der Tonfolge als

Wahrnehmungsgegenstand sich aufzuheben beginnt.

Da die Variabilität des Merkmals "Tonhöhe" aber nicht nur von $\Delta f_{1,2}$ sondern auch von $\Delta f_{2,3}$ abhängt, sollte auch die dreifache Interaktion zwischen den Faktoren $\Delta f_{1,2}$, $\Delta f_{2,3}$ und Intervallkonfiguration signifikant sein und eine um so g r ö s s e r e RT in Abhängigkeit von $\Delta f_{1,2}$ zeigen, je k l e i n e r $\Delta f_{2,3}$ ist. Wie Abb. 7 zeigt, ist dies auch tatsächlich der Fall ($F(18, 522) = 4.05$, $p<0.01$; bei df = 1,29 $p<0.10$).

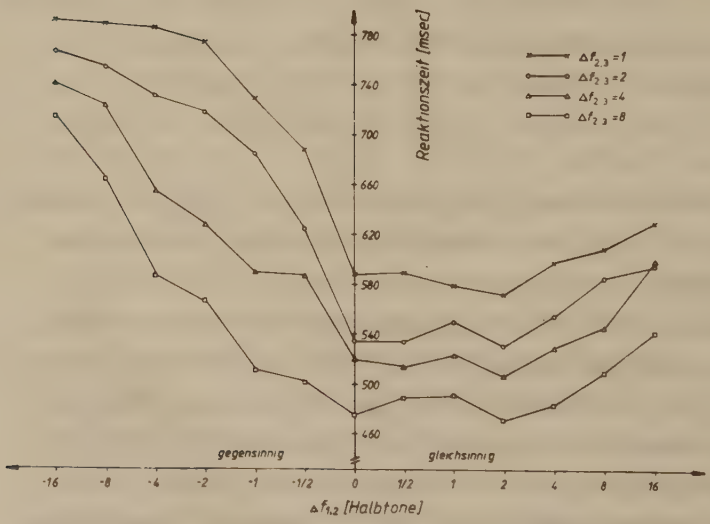


Abb. 7: RT-Mittelwerte für die Interaktion " $\Delta f_{1,2}$ x $\Delta f_{2,3}$ x Intervallkonfiguration"

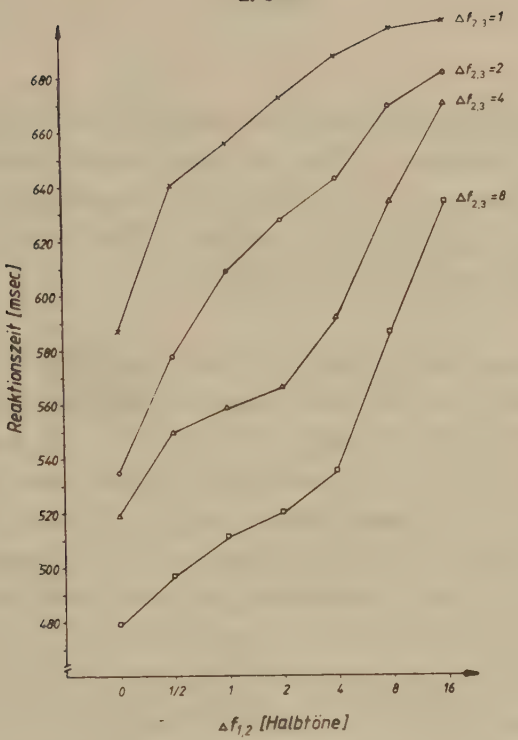


Abb. 8: RT-Mittelwerte für die Interaktion " $\Delta f_{1,2}$ X $\Delta f_{2,3}$ "

Bei kleinem $\Delta f_{2,3}$ kann der Verlust an Objektpermanenz bereits bei mittlerem $\Delta f_{1,2}$ auftreten, so daß die RT bei weiterer Zunahme von $\Delta f_{1,2}$ nur noch geringfügig oder aber gar nicht steigen sollte, aber bei großem $\Delta f_{2,3}$ sollte der Punkt des Verlustes an Objektpermanenz erst bei weitaus größeren Werten von $\Delta f_{1,2}$ eintreten, so daß die RT in diesem Fall über einen weiteren Bereich von $\Delta f_{1,2}$ zunehmen sollte. Diese Steigungsunterschiede sind bei gegensinnigem Konturverlauf in Abb. 6 deutlich zu sehen, während sie bei gleichsinnigem Konturverlauf nicht zu beobachten sind.

Diese Steigungsunterschiede sind bei gegensinnigem Konturverlauf aber so stark ausgeprägt, daß sie sich auch in der Interaktion zwischen den Faktoren $\Delta f_{1,2}$ und $\Delta f_{2,3}$ noch unabhängig von der Intervallkonfiguration nachweisen lassen (s. Abb. 8; $F(18, 522) = 3.52$, $p \leq 0.01$; bei df = 1,29 $p \leq 0.10$).

Gleichzeitig zeigt Abb. 8, daß es offenbar noch einen $\Delta f_{2,3}$ - Haupteffekt geben muß. Die RT-Mittelwerte dieses Haupteffekts zeigt Abb. 9 ($F(3,87) = 29.49$, $p \leq 0.01$; bei df = 1,29 ebenfalls $p \leq 0.01$). Der $\Delta f_{2,3}$ - Haupteffekt verläuft genau umgekehrt wie der $\Delta f_{1,2}$ - Haupteffekt (s.Abb. 6). Offenbar erleichtert ein größeres $\Delta f_{2,3}$ die Entnahme der aufgabenrelevanten Information über seine Verlaufsrichtung erheblich, indem die Diskriminationsschwierigkeiten bei der Entscheidung zwischen auf- und absteigender Kontur, die bei kleinen Intervallen gewiß gegeben sind, bei großen Intervallen reduziert werden (vgl. Divenyi & Hirsh, 1974; Nickerson & Freeman (1974); Ortmann (1926); Warren (1972); Warren & Byrnes (1975)).

Das Diskriminationsproblem bei der Entnahme der aufgabenrelevanten Information sollte sich aber nur dann einstellen, wenn von dem Niveau der Gesamtrepräsentation auf das unmittelbar darunterliegende Niveau zum Zwecke der Fokussierung des letzten Intervalls gewechselt werden muß. Dies ist aber nur bei gegensinniger Intervallkonfiguration der Fall, so daß die Reduktion der RT bei großem $\Delta f_{2,3}$ sich bei gegensinnigem Verlauf stärker als bei gleichsinnigem zeigen sollte. Diesen Steigungsunterschied zeigt Abb. 1o ($F(3,87) = 2o.88$, $p \leq 0.01$; bei df = 1,29 ebenfalls $p \leq 0.01$).

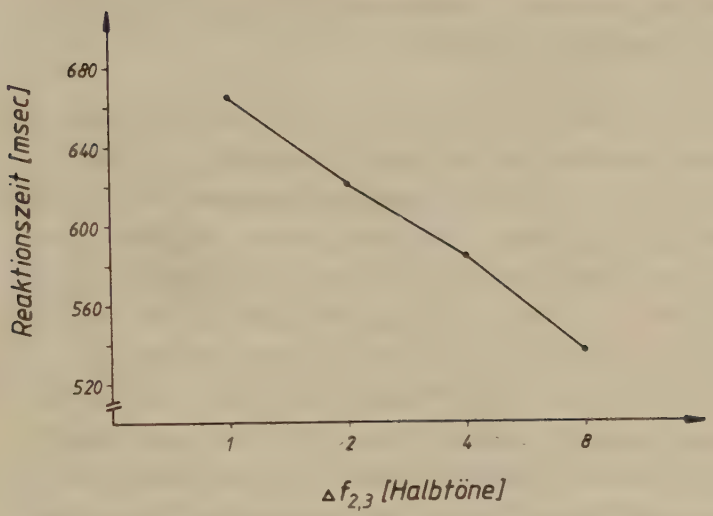


Abb. 9: RT-Mittelwerte für $\Delta f_{2,3}$ - Haupteffekt

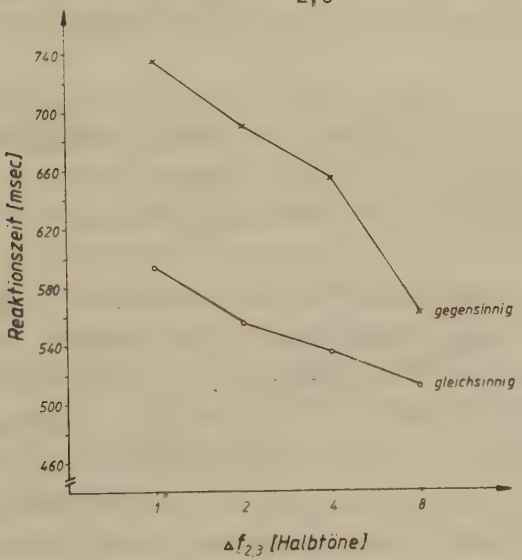


Abb. 10: RT-Mittelwerte für die Interaktion " $\Delta f_{2,3}$ X Intervallkonfiguration"

Der Einfluß der zeitlichen Gruppierung der Töne auf die Fokussierbarkeit des letzten Intervalls

Die Variation von $SOA_{1,2}$ und $SOA_{2,3}$ sollte, so unsere Hypothese, Bedingungen herstellen, die bei relativ großem $SOA_{1,2}$ im Vergleich zum $SOA_{2,3}$ ein auditives Chunking begünstigen, das die beiden letzten Töne als eigene Verarbeitungseinheit vom ersten Ton isoliert, während bei relativ kleinem $SOA_{1,2}$ im Vergleich zum $SOA_{2,3}$ ein Chunk entstehen sollte, das die beiden ersten Töne vom letzten isoliert. Wenn die aufgabenrelevante Information nicht der Gesamtrepräsentanz entnommen werden kann, wie dies bei gegensinniger Intervallkonfiguration der Fall ist, sollte eine Isolation der beiden letzten Töne dazu führen, daß diese direkt fokussiert werden können, während bei Isolation des letzten Tons die Entnahme der aufgabenrelevanten Information nur über das dem Chunkniveau untergeordnete Elementniveau möglich sein sollte. In letzterem Fall müßte daher die RT länger als im ersteren sein.

Da diese Hypothese das Verhältnis der beiden SOAs und die Intervallkonfiguration betrifft, müssen wir die dreifache Interaktion dieser Faktoren betrachten, die in Abb. 11 dargestellt ist ($F(3,87) = 7.95$ $p \leq 0.01$; bei df = 1,29 ebenfalls $p \leq 0.01$).

Wie wir Abb. 11 entnehmen können, ist der Kurvenverlauf für gegensinnige und gleichsinnige Intervallkonfiguration unterschiedlich: Die erwartete Abnahme der RT bei größerem $SOA_{1,2}$ und gegensinniger Intervallkonfiguration deutet sich bei beiden $SOA_{2,3}$ - Bedingungen an (bei $SOA_{2,3}$ = 18o msec ein Abfall von insgesamt 3o msec, bei $SOA_{2,3}$ = 36o msec von nur 2o msec), aber bei gleichsinniger Intervallkonfiguration finden wir einen U-förmigen Verlauf, dessen RT-Minimum immer dort liegt, wo $SOA_{1,2} = SOA_{2,3}$ ist.

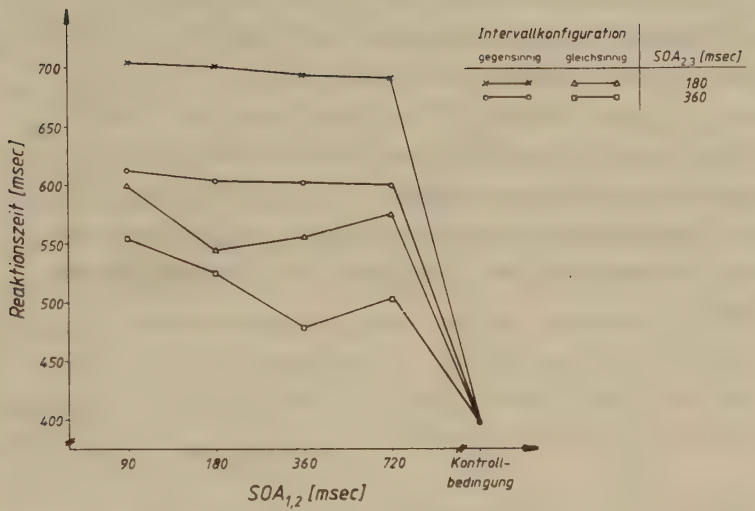


Abb. 11: RT-Mittelwerte der Interaktion "$SOA_{1,2}$ X $SOA_{2,3}$ X Intervallkonfiguration"

Die Abnahme der RT im $SOA_{1,2}$ - Haupteffekt (s.Abb. 3, S. 164) zwischen 90 und 180 msec, die die einzige signifikante Differenz sukzessiver $SOA_{1,2}$ - Bedingungen ist (Scheffé-Test, $p \leq 0.01$), geht demnach überwiegend auf die Bedingungen mit gleichsinniger Intervallkonfiguration zurück, und das Ausbleiben signifikanter sukzessiver Paardifferenzen im $SOA_{1,2}$ - Haupteffekt geht ebenfalls auf den Anstieg der RT bei $SOA_{1,2} \geq SOA_{2,3}$ und gleichsinniger Intervallkonfiguration zurück, der den Abfall bei gegensinnigem Verlauf in den Haupteffektmittelwerten kompensiert.

Den unterschiedlichen Effekt von $SOA_{1,2}$ auf die RT beider Intervallkonfigurationen zeigt auch die Interaktion "$SOA_{1,2}$ X Intervallkonfiguration" (s.Abb. 12; $F(3,87) = 9.24$, $p \leq 0.01$; bei $df = 1,29$ ebenfalls $p \leq 0.01$). Hier erweist sich der RT-Abfall zwischen 9o und 72o msec $SOA_{1,2}$ bei gegensinniger Intervallkonfiguration als nicht signifikant (Scheffé-Kontraste, $p > 0.10$), während der zwischen 9o und 36o msec $SOA_{1,2}$ bei gleichsinniger Intervallkonfiguration sehr signifikant ist (Scheffé-Kontraste, $p \leq 0.01$).

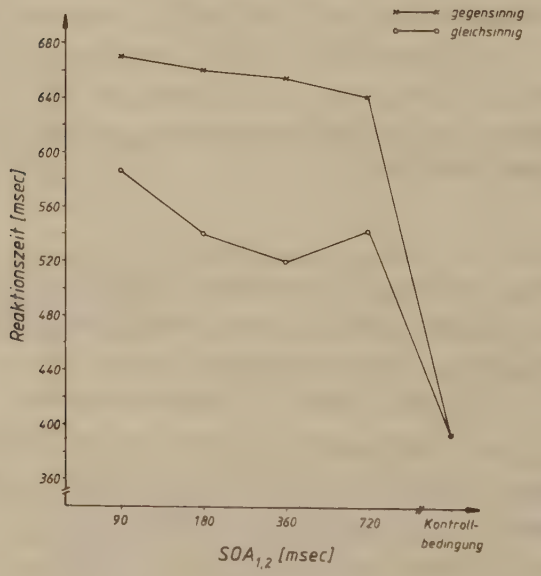


Abb. 12: RT-Mittelwerte für die Interaktion "$SOA_{1,2}$ X Intervallkonfiguration"

Entgegen unserer Hypothese wirkt sich ein relativ großes $SOA_{1,2}$ bei gegensinnigem Verlauf kaum reaktionszeitmindernd aus. Dagegen haben wir bei gleichsinnigem Verlauf ein Befundmuster, das wir auf r h y t h m i s c h e H o m o g e n i t ä t der Tonfolgen bei $SOA_{1,2} = SOA_{2,3}$ zurückführen können. Bei rhythmischer Homogenität kommt zur Invarianz der Intervallrichtungen noch die Invarianz der zeitlichen Abstände zwischen den Tönen hinzu. Ein solches Muster läßt bereits auf dem obersten Repräsentationsniveau eindeutig erkennen, daß die aufgabenrelevante Information diesem zu entnehmen ist, während bei rhythmischer Inhomogenität offenbar noch eine detailliertere Analyse der auf dem obersten Niveau verfügbaren Informationen erforderlich ist.

Wenn die eben formulierte zweite Erklärung des $\Delta f_{1,2}$ - Effekts zutrifft, die besagt, daß bei Verlust der Objektpermanenz wegen zu geringer Invarianz des Merkmals "Tonhöhe" die Gesamtrepräsentanz keine konfigurativen Merkmale enthält, die sich auf alle drei Töne erstreckt, und deshalb die Steuerung der Fokussierung erschwert und dadurch verlängert wird, dann sollte eine besonders deutliche Repräsentation des letzten Intervalls bei großem $SOA_{2,3}$ die Fokussierung und Entnahme der Intervallrichtungsinformation zumindest teilweise wieder erleichtern, d.h. die RT-Differenz zwischen den $SOA_{2,3}$ - Bedingungen bei 180 msec SOA und 360 msec SOA sollte bei größer werdendem $\Delta f_{1,2}$ zunehmen. Abb. 13 zeigt, daß diese Interaktion zwischen $\Delta f_{1,2}$ und $SOA_{2,3}$ tatsächlich in dieser Richtung vorhanden ist ($F(6,174) = 5.07$, $p<0.01$; bei df = 1,29 $p < 0.05$).

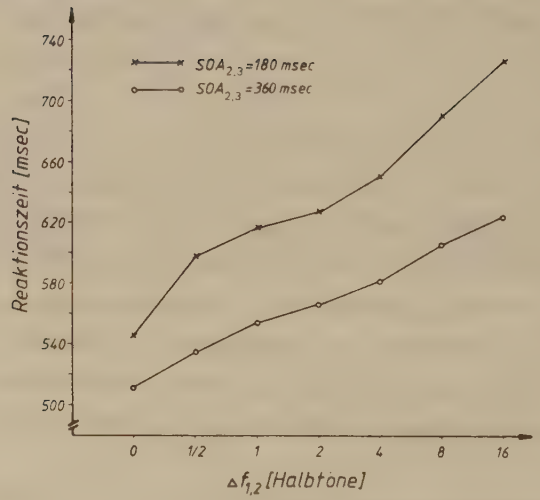


Abb. 13: RT-Mittelwerte für die Interaktion " $\Delta f_{1,2}$ X $SOA_{2,3}$"

Abhängigkeit der Konturdiskrimination vom SOA zwischen den beiden letzten Tönen

Zum Abschluß berichten wir noch die verbleibenden signifikanten Effekte, zu denen wir keine Hypothesen formuliert haben. Ihnen gemeinsam ist jeweils die Beteiligung des Faktors $SOA_{2,3}$ an der Variation der RT-Daten. Der $SOA_{2,3}$ - Haupteffekt ist sehr signifikant ($F(1,29) = 24.91$, $p \leq 0.01$) und zeigt eine Abnahme der RT bei $SOA_{2,3}$ = 36o msec (Mittelwert: 569 msec) gegenüber dem $SOA_{2,3}$ = 18o msec (Mittelwert: 636 msec).

Wenn dieser Effekt Ausdruck einer leichteren Konturdiskrimination bei großen im Vergleich zu kleinen SOAs ist, dann sollte eine stärkere Abnahme der RT zu beobachten sein, wenn das letzte Intervall fokussiert werden muß. Dies ist bei gegensinnigem, aber nicht bei gleichsinnigem Konturverlauf der Fall, so daß die RT bei gegensinnigem Konturverlauf von der Variation des $SOA_{2,3}$ stärker als bei gleichsinnigem Verlauf beeinflußt werden sollte.

Die Mittelwerte der Interaktion "$SOA_{2,3}$ X Intervallkonfiguration" bestätigen diese Schlußfolgerungen (s.Abb. 14; $F(1,29) = 15.37$, $p=0,01$): Die Abnahme der RT für gegensinnige Konfigurationen ist steiler als die Abnahme für gleichsinnige Konfigurationen.

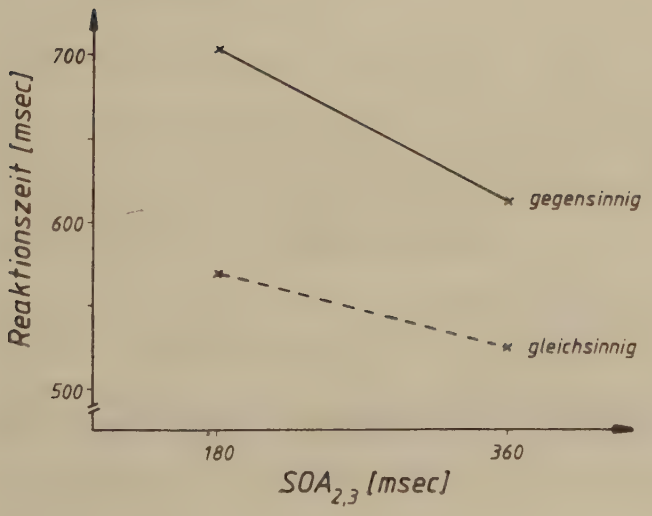


Abb. 14: RT-Mittelwerte für die Interaktion "$SOA_{2,3}$ X Intervallkonfiguration"

Zur Diskriminierbarkeit zwischen den Richtungen der letzten Intervalle trägt aber nicht nur das SOA zwischen den Tönen bei, sondern auch die Intervallgröße. Die Interaktion beider Faktoren zeigt Abb. 15 ($F(3,87) = 3.92$; $p \leq 0.02$; bei $df = 1,29$ $p \leq 0.10$).

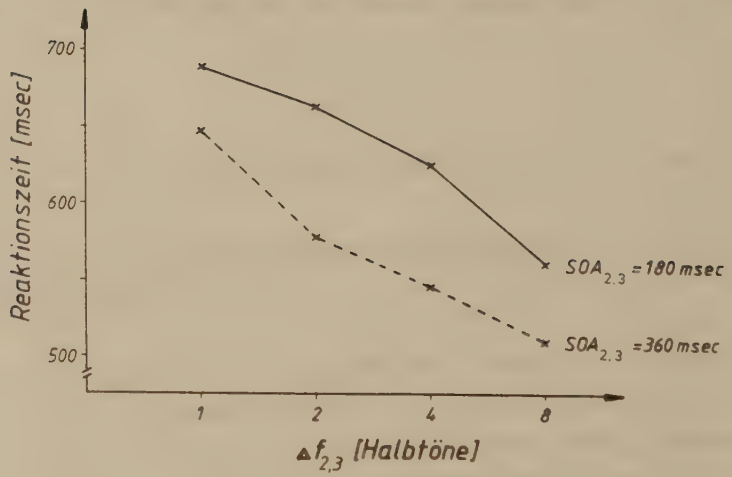


Abb. 15: RT-Mittelwerte für Interaktion " $\Delta f_{2,3}$ X $SOA_{2,3}$ "

Die Kurve für $SOA_{2,3}$ = 18o msec verläuft positiv beschleunigt, die für $SOA_{2,3}$ = 36o msec negativ beschleunigt, so daß sich die Kurven in zwei Punkten schneiden würden, wenn man sie in beiden Richtungen extrapolieren würde. Diese Datenkonfiguration läßt vermuten, daß es eine Grenze für die Förderung der Diskriminierbarkeit gibt: Bei $\Delta f_{2,3}$ in der Nähe von Null, also bei einem nicht mehr diskriminierbaren Frequenzunterschied zwischen beiden Tönen, wäre ein auch noch so großer zeitlicher Abstand kaum mehr für die Diskrimination förderlich; bei großen Werten von $\Delta f_{2,3}$ sollte die Diskriminierbarkeit auch durch eine weitere Zunahme von $SOA_{2,3}$ kaum noch weiter zu verbessern sein, sondern im Gegenteil bei extrem großen Werten von $SOA_{2,3}$ aufgrund eines Verlustes an Kohärenz zwischen beiden Tönen wieder verschlechtert werden.

5.4 ZUSAMMENFASSUNG

Experiment 1 thematisierte zwei theoretisch zentrale Fragen der auditiven Mustererkennung, die sich aus der referierten Befundlage ergeben hatten und die wir auf dem Hintergrund des Hierarchiemodells unter Einbeziehung des Konzepts des auditiven Chunkings glaubten theoretisch beantworten zu können.

Die erste Frage bezog sich auf die Entscheidung zwischen einem Modell der kognitiven Verarbeitung von Strukturen in Tonfolgen, das eine von den tonalen Elementen ausgehende Kodierung einer Repräsentation der Gesamttonfolge annimmt (lineares Informationsverarbeitungsmodell), und einem Modell, in dem die Verarbeitung

von der Gesamtrepräsentation einer Tonfolge ausgeht und die Analyse und Entnahme der aufgabenrelevanten Information als ein Durchlaufen einer hierarchisch gegliederten Repräsentationsstruktur beschreibt, in der die aufgabenrelevante Information auf dem jeweils höchsten Niveau entnommen wird, das sie enthält (Hierarchiemodell).

Die zweite Frage zielte auf die funktionalen und strukturellen Ursachen der differentiellen Schwierigkeiten des Erkennens tonaler Strukturen in Abhängigkeit von der Art der Konfiguration aller Töne. Die theoretische Erklärung hierfür im Rahmen des Hierarchiemodells führte die unterschiedlichen Schwierigkeiten auf die Anzahl der zu durchlaufenden und zu analysierenden strukturellen Repräsentationsniveaus der Tonfolge zurück, die in Abhängigkeit von der Konfiguration der Töne variieren sollte, und auf ein von der Variation der physischen Merkmale der Tonfolge abhängiges auditives Chunking, das ein mittleres Niveau von Verarbeitungseinheiten zwischen dem der Gesamtrepräsentation und dem der Elementarrepräsentanzen konstituiert.

Das Experiment bestand in einer Zwei-Alternativen-Wahlreaktionsaufgabe, in der die Intervallrichtung der beiden letzten Töne einer Dreitonfolge zu identifizieren war. Abhängige Variable war die Reaktionszeit; als unabhängige Variablen wurden die Intervallgrößen zwischen den drei Tönen, die SOAs zwischen ihnen und die Gesamtkonfiguration aller drei Töne variiert.

Bezüglich der ersten Frage unterstützt das Befundmuster eindeutig das Hierarchiemodell: Die RT für Tonfolgen, deren Gesamtrepräsen-

tation bereits die aufgabenrelevante Information enthält, erwies sich als bedeutend kürzer als die für Tonfolgen, deren Gesamtrepräsentation die aufgabenrelevante Information nicht enthält und deshalb einen Übergang auf ein niedrigeres Repräsentationsniveau notwendig macht, aus dem diese Information zu entnehmen ist.

Das für die Beantwortung der zweiten Frage relevante Befundmuster war dagegen weniger klar. Die Variationen physischer Merkmale, die ein auditives Chunking bewirken sollten und dadurch unter bestimmten Bedingungen das kritische letzte Intervall vom ersten nicht zu beachtenden Ton als eigene Verarbeitungseinheit isolieren sollten, erwiesen sich nicht in der vorhergesagten Richtung als wirksam: Ein großes Intervall zwischen erstem und zweitem Ton führte nicht zu der erwarteten Chunkbildung zwischen zweitem und drittem Ton und einer hierdurch leichteren Fokussierbarkeit des aufgabenrelevanten strukturellen Details der Tonfolge, sondern erhöhte entgegen unseren Erwartungen die RT so stark, daß eine Erklärung dieses Effekts durch Chunking nicht möglich war. Außerdem führte die Variation der SOAs zwischen Tönen nicht zu der aufgrund der Chunkingkonzeption zu erwartenden signifikanten Reduktion der Reaktionszeit, obwohl bei gegensinniger Intervallkonfiguration diese Reduktion der Tendenz nach eintrat.

Wir konnten durch ad hoc Annahmen allerdings das Hierarchiemodell "retten", indem wir davon ausgingen, daß das Ausbleiben des erwarteten Chunkingeffekts durch einen Verlust an Objektpermanenz bei sehr großem ersten Intervall aufgrund zu geringer Invarianz der physischen Merkmale über die Zeit bewirkt wird,

die zur Folge hat, daß die Gesamtrepräsentation nicht mehr die konfigurativen Merkmale der Tonfolge enthält, sondern nur noch deren zeitliche Abfolge als Sukzession. Die hierdurch erschwerte Steuerung der Fokussierung des letzten Intervalls führt dann zu der beobachteten Erhöhung der RT.

Dieses Befundmuster macht ein Überdenken der Chunkingkonzeption erforderlich sowie eine theoretische Integration des Gibson'schen Invarianzkonzepts in das Hierarchiemodell. Dieses wird im folgenden Kapitel zu leisten und durch ein weiteres Experiment dann zu überprüfen sein.

6. EXPERIMENT 2: WAHRNEHMUNG UND IDENTIFIKATION STRUKTURELLER INVARIANZEN IN TONFOLGEN

6.1 EINLEITUNG

Fassen wir die Befunde von Experiment 1 zusammen, dann zeigt sich sowohl in den erwarteten als auch in den erwartungsdiskrepanten Befunden ein gemeinsames Prinzip der kognitiven Verarbeitung und Repräsentation von Tonfolgen: Eine Veränderung relativ lokaler Merkmale (Tonhöhe) oder Merkmalsrelationen (Intervallgröße und -richtung, zeitliche Abstände zwischen Tonhöhen) bewirkt eine Veränderung der globalsten Repräsentation der Tonfolge und determiniert über diese die Schwierigkeit der Entnahme aufgabenrelevanter Informationen.

Dieses Prinzip erweist sich sowohl bei Betrachtung des erwarteten Haupteffekts der Intervallkonfiguration als zutreffend als auch bei einer Analyse der unerwarteten Befundmuster zum Effekt von $\Delta f_{1,2}$ und $SOA_{1,2}$: Die $SOA_{1,2}$ - Variation führte zu Reaktionszeitdifferenzen zwischen Bedingungen, wie sie aufgrund der Chunking-Konzeption nicht zu erwarten gewesen wären; hier zeigte sich ein Effekt, der bei minimaler SOA-Variation ($SOA_{1,2} = SOA_{2,3}$) eine besonders leichte Identifikation des letzten Intervalls ermöglichte; und auch der $\Delta f_{1,2}$ - Effekt entsprach nicht den Vorhersagen auf der Basis der Chunking-Konzeption; vielmehr zeigte er komplexe Wechselwirkungen, die deutlich machen, daß bei minimaler Variation der Intervallgröße und -richtungen zwischen allen Tönen das letzte Intervall am leichtesten zu identifizieren ist.

Aus diesen Befunden müssen wir die beiden folgenden theoretischen Konsequenzen ziehen:

1. Auf den einzelnen Niveaus der Repräsentationshierarchie sind nicht Konfigurationen e i n z e l n e r Merkmale, sondern nur Konfigurationen von M e r k m a l s b ü n d e l n zu entnehmen, in die einzelne Merkmale als integrale Bestandteile eingehen.

2. Die Analyse eines tonalen Verlaufsmusters besteht nicht in der Verarbeitung diskreter, statischer Einheiten, sondern vielmehr in der Verarbeitung der V a r i a t i o n s s t r u k t u r e i n e s M e r k m a l s b ü n d e l s in bezug auf die Zeit.

Akzeptieren wir diese beiden Aussagen als Prämissen eines Modells der kognitiven Verarbeitung von Tonfolgen und Musik, dann müssen wir das Konzept des Chunkings bzw. das diskreter Verarbeitungseinheiten aufgeben und stattdessen annehmen, daß die Form der Repräsentation eines sich dynamisch über die Zeit entwickelnden Reizmusters eine a n a l o g e ist, zu der erst bei kategorialer Abbildung der Verlaufsstruktur eines Repräsentationsniveaus eine diskrete Form der Repräsentation hinzukommt.

Chunks und Verarbeitungseinheiten sind theoretische Konstrukte, die ihre Herkunft der Computeranalogie der Informationsverarbeitung zu verdanken haben. Sie wurden zur Erklärung von Phänomenen verwendet, wie sie nur in einer paradigmatischen experimentellen Situation, in der mit statischer Reizdarbietung gearbeitet wird, auftreten können. Wahrnehmung vollzieht sich dagegen in einer sich

ständig verändernden Umwelt, die phänomenal kaum etwas Statisches an sich hat. Eine statische Konzeption der Informationsverarbeitung stößt da an die Grenzen ihres Erklärungswertes, wo es gilt, Wahrnehmung außerhalb der restriktiven Bedingungen des Labors zu erklären (vgl. Neisser, 1976).

Daß auch bei der Identifikation dynamischer, zeitlicher Reizmuster an einer Stelle des Informationsverarbeitungsprozesses eine "Digitalisierung" der Reizmusterrepräsentation notwendig ist, soll hier nicht bestritten werden. Auch die theoretische Annahme von Wolff (1977), daß das Reizmuster bereits vor einer Entnahme von Identitätsinformationen in strukturierter Form repräsentiert ist, soll nicht aufgegeben werden. Die theoretischen Probleme, die sich uns jetzt neu stellen, bestehen 1. in der Fixierung des Ortes einer "Digitalisierung" der repräsentierten Information über ein Reizmuster und 2. in der Beschreibung eines hierarchisch strukturierten, analogen Repräsentationssystems.

In der Konzeption von Wolff (1977) besteht die Repräsentationshierarchie aus abgegrenzten Einheiten, die als Merkmalsverband bezeichnet werden, denen genau ein Ort zugeordnet ist. Will man diese statische Betrachtung der Repräsentationshierarchie in eine dynamische überführen, braucht man lediglich noch die Dimensionen als einheitsstiftende Variablen hinzuzufügen, deren Variation die Einheit des Merkmalsverbandes unangetastet lassen. Für die visuelle Modalität sind dies Raum und Zeit, d.h. eine potentiell fokussierbare Repräsentation liegt dann vor, wenn der Merkmalsverband bei Variation in Raum und Zeit invariant bleibt. Analog dazu ist eine auditive Repräsentation dann poten-

tiell fokussierbar, wenn die Variation der Merkmale eines akustischen Reizmusters über die Zeit invariant bleibt. Die "Digitalisierung", d.h. die Bildung statischer Repräsentationen auf der Grundlage einer analogen Repräsentation eines tonalen Verlaufsmusters, kann in diesem Fall erst als Ergebnis eines Entnahme- und Identifikationsprozesses gesehen werden, dem Identitätskategorien für Verlaufsstrukturen zur Verfügung stehen.

Die analoge Form der Repräsentationshierarchie als Struktur der Informationen des echoischen Speichers stellt eine z e i t l i c h r e l a t i v i e r t e Repräsentation von Merkmalsvariationen dar; die aufgrund von Identifikationsprozessen repräsentierte Verlaufsstruktur dagegen stellt eine n i c h t z e i t l i c h e Repräsentation der Merkmalsvariationen dar, aus der zeitliche Relationen aus der Organisationsstruktur der Repräsentation (vertikale und horizontale Struktur) erschlossen werden können.

Das Invarianzkonzept Gibson's

Akustische Signalfolgen sind immer das Resultat mehr oder weniger systematischer physischer Veränderungen in der Umwelt. Die Formen dieser Veränderungen reichen von der mikroskopischen Korngröße der Variation des Schalldrucks über die Zeit bis zu makroskopischen Variationen, wie sie uns als Verarbeitung eines Themas in der Durchführung eines Sonatenkopfsatzes begegnen.

Gibson's Theorie der "direkten Wahrnehmung" (Gibson, 195o, 1966a, b, 1977) geht davon aus, daß Wahrnehmen nicht in der Integration

statischer Schnappschüsse bestehe, sondern daß Wahrnehmen in der Repräsentation dessen besteht, was bei Änderung eines Teils der physischen Merkmale von Objekten und ihrer Umgebung in Raum und Zeit konstant, oder wie er sagt, invariant bleibt (vgl. auch Bransford, McCarrell, Franks & Nitsch, 1977; Mace, 1974, 1977; Turvey, 1975, 1977; Shaw & Bransford, 1977; Shaw, McIntyre & Mace, 1974; Shaw & Pittenger, 1978). Die Wahrnehmung einer bestimmten Form von Transformation, der die Merkmalsbündel eines Objekts bei Veränderung in Raum und Zeit unterliegen, bedingt seine Permanenz als Wahrnehmungsgegenstand auf einem Hintergrund von Merkmalsänderungen, die nicht derselben Form von Transformation gehorchen. Wahrnehmung besteht demnach im Entdecken von S t i l e n d e r V e r ä n d e r u n g in der physischen Welt, oder in Gibson's Terminologie, im Entdecken von Invarianzen (zum Invarianzkonzept vgl. auch Carr & Bacharach, 1976; Mace, 1974; Turvey, 1975).

Den Vorgang des Entdeckens von Invarianzen bezeichnet Gibson (1966) als "tuning" durch "Resonanz". Shaw & McIntyre (1974) vergleichen dieses "tuning" mit dem Abstimmen eines Senders beim Einstellen eines Radios: Wahrnehmung wird dadurch möglich, daß das kognitive System auf die "Wellenlänge" oder "Frequenz" der Variation einer Stimulation abgestimmt wird. Wenn der Organismus auf eine bestimmte Form der Merkmalsvariation abgestimmt ist, "a match occurs and a pattern of 'oscilliation' is set up in the organism which corresponds to some specific pattern of oscilliation in the enviroment" (Carr & Bacharach, 1976, S. 29o).

Invarianzen lassen sich hierarchisch nach ihrer Ordnung gliedern: Die Ordnung einer Invarianz ist durch die Anzahl gleichzeitig variierender Merkmalsdimensionen gegeben, die der Invarianz zugrunde liegen (Shaw & McIntyre, 1974). Auf den Fall der Dreitonfolgen im ersten Experiment angewandt, heißt dies, daß bei alleiniger Variation des Merkmals Tonhöhe eine Invarianz erster Ordnung vorliegt (gleichsinnige Intervallkonfiguration mit konstanten Intervallgrößen und SOAs); variiert zusätzlich noch die Richtung der Intervalle (gegensinnige Intervallkonfiguration mit konstanten Intervallgrößen bei konstanten SOAs), liegt eine Invarianz zweiter Ordnung vor, und variiert außerdem noch das SOA zwischen den Tönen, dann handelt es sich um eine Invarianz dritter Ordnung (gegensinnige Intervallkonfiguration bei $SOA_{1,2} \neq SOA_{2,3}$).

Diese Invarianzordnung der Struktur von Dreitonfolgen reflektiert ziemlich exakt die Schwierigkeiten, die die Vpn beim Erkennen der Richtung des letzten Intervalls in Experiment 1 hatten: je höher die Invarianzordnung desto länger die RT.

Repräsentation von Invarianzstrukturen

Eine Invarianzstruktur wird als Hierarchie von Invarianzen repräsentiert, die in vertikaler Richtung von unten nach oben aufsteigend nach ihrer Ordnung ineinander zeitlich verschachtelt sind. Auf einer Ebene der vertikalen Schichtung folgen Invarianzen gleicher Ordnung in zeitlicher Sukzession. Invarianzen gleicher Ordnung können verschiedenen Typs sein: Ein Ausschnitt der hori-

zontalen Gliederung eines Repräsentationsniveaus kann z.B. in einer Invarianz hinsichtlich des Merkmals Tonhöhe bestehen, während das Merkmal Pausenlänge zwischen Tönen variiert, und der unmittelbar nachfolgende Ausschnitt kann hinsichtlich des Merkmals Pausenlänge invariant sein, aber eine Variation des Merkmals Tonhöhe enthalten. Die Ordnung zweier in vertikaler Richtung benachbarter Invarianzen ist um mindestens einen Ordnungsgrad unterschiedlich.

Sind alle physischen Merkmale einer Tonfolge konstant, liegt eine Invarianz nullter Ordnung vor. Variiert ausschließlich das Merkmal Tonhöhe in konstanten Intervallen, konstituiert dessen Variation eine Invarianz erster Ordnung, nämlich die Intervalle gleicher Größe. Kommt nun noch die Variation der Intervallgrößen hinzu, dann liegt Invarianz zweiter Ordnung vor, nämlich konstante Intervallrichtungen oder eine konstante Kontur. Variiert darüber hinaus auch noch die Intervallrichtung, werden Invarianzen dritter Ordnung erzeugt, die verschiedene Typen von Konturverläufen mit Richtungswechseln darstellen. Diese verschiedenen Konturverlaufstypen können in ihrer zeitlichen Reihenfolge wiederum invariant sein oder eine Variation aufweisen, die verschiedene Komplexitätsgrade annehmen kann. So stellt z.B. ein symmetrischer Verlauf (ein Takt mit aufsteigender Konturlinie, der folgende mit absteigender) eine Invarianz höherer Ordnung dar als eine Wiederholung. Eine sehr hohe Invarianz ist z.B. die Konstanz der Periodengliederung eines Volksliedes bei Vergleich mehrerer Volkslieder. Obwohl sämtliche physischen Merkmale (Lautstärke, Klangfarbe, Akzentverteilung, Tonhöhe usw.) ganz andere Variationen zeigen,

bleiben diese aber insofern auf einem Niveau sehr großer Korngröße strukturell invariant, als die Form ihrer zeitlichen Verschachtelung, also die vertikale Gliederung, dieselbe bleibt: Es schließen sich immer Paare horizontal benachbarter Invarianztypen zu einer Invarianz höherer Ordnung zusammen.

Der Rhythmus ist in den meisten Fällen eine Invarianz der fünften Ordnung: Es variieren gleichzeitig die Tonhöhe, die Pausenlänge, die Intervallgröße und -richtung, bevor sich die Akzentstärke eines Tones wiederholt. Da es aber gerade der Rhythmus ist, der bei deutlicher Ausprägung das Verfolgen von Musik so stark leitet (z.B. Dowling, 1973a), dieser aber bereits eine Invarianz mittlerer Ordnung darstellt, liegt die Annahme nahe - die auch durch die alltägliche Hörerfahrung bestätigt wird -, daß beim Musikhören primär Invarianzen mittlerer Ordnung wahrgenommen werden.

Diese mittlere Invarianzordnung besitzt eine Periode (Abstand zwischen identischer Wiederholung der den Rhythmus konstituierenden Merkmalsvariationen), die die psychische Präsenzzeit nicht überschreitet (Fraisse, 1978). Aus diesem Grunde sind innerhalb der Zeitspanne der psychischen Präsenzzeit auch Invarianzen niedrigerer Ordnung fokussierbar, also strukturelle Detailanalysen niedrigerer Korngröße möglich, während in dem Fall, in dem die zeitliche Periode einer Invarianz höherer Ordnung größer als die zeitliche Erstreckung der psychischen Präsenzzeit ist, nur noch solche Invarianzen direkt fokussierbar sind, deren Periode sich mit der Erstreckung der psychischen Präsenzzeit überlappt, d.h. ein Invarianztyp niedrigerer Ordnung, der bereits weiter

zurückliegt, kann nur über die Invarianz höherer Ordnung erschlossen werden, die sich innerhalb der psychischen Präsenzzeit konstituiert. Z.B. kann die Invarianzbeziehung zwischen den Takten eines Motivs dann unmittelbar fokussiert werden, wenn sich diese Invarianz innerhalb der psychischen Präsenzzeit gerade konstituiert; liegt dieses Taktpaar aber bereits eine Weile zurück, dann kann seine Beziehung nur über die höhere Invarianzordnung entnommen werden, die als Invarianz niedriger Ordnung dieses Taktpaar enthält, deren Periode aber noch innerhalb der psychischen Präsenzzeit liegt. Dies ist z.B. am Ende einer viertaktigen Phrase der Fall, deren zwei letzte Takte noch innerhalb der psychischen Präsenzzeit liegen, während die beiden ersten Takte bereits außerhalb liegen. In diesem Fall muß die Beziehung der beiden ersten Takte auf dem Niveau der Phrase als die ihnen übergeordnete Invarianz entnommen werden.

Das hat also zur Folge, daß die Detailliertheit einer Repräsentation eines Ausschnitts um so geringer wird, je länger dieser Ausschnitt zeitlich zurückliegt, weil das Niveau, über das diese Repräsentation verfügbar wird, mit zunehmendem zeitlichen Abstand von den psychisch präsenten Repräsentationen immer globaler wird.

"Tuning" als Fourieranalyse

Um auf der Grundlage des Gibson'schen Invarianzkonzepts zu einem funktionalen Modell der Strukturwahrnehmung in Tonfolgen und Musik

zu gelangen, müssen wir den Prozeß des "tuning" (im folgenden "Abstimmung" genannt) funktional näher spezifizieren, als dies Gibson selbst getan hat.

Abstimmung heißt, ein Repräsentationsniveau dadurch fokussierbar zu machen, daß sich das kognitive System auf eine bestimmte zeitliche Periode der Variabilität der Merkmale und damit auf eine bestimmte Invarianzordnung einstellt. Da wir davon ausgegangen sind, daß nicht die Variationen einzelner Merkmalsdimensionen, sondern die eines ganzen Merkmalsverbands repräsentiert werden, kann das Abstimmen im Prinzip als eine Art "makroskopischer Fourieranalyse" der mehrdimensionalen Verteilung aller Merkmale einer bestimmten Invarianzordnung betrachtet werden.

Wir nehmen an, daß kognitiv der Verlauf einer mehrdimensionalen, zeitlichen Merkmalsverteilungsfunktion repräsentiert ist, die durch Superposition der Verteilungen einzelner Merkmale zustandekommt. Die Aufgabe des kognitiven Systems besteht nun darin, die Dimensionalität dieser Funktion so zu reduzieren, daß invariante Abschnitte niedriger Ordnung nicht durch Verteilungen höherer Ordnung "maskiert" werden. Auf diese Weise sollten Invarianzen entdeckbar sein.

Wenn die Repräsentation der kognitiven Verteilungsfunktion Inhalt des echoischen Speichers ist, dann muß der Abstimmungsmechanismus einer sein, der in der strukturellen Abfolge der Verarbeitungsschritte dem Zugriff auf die Information des echoischen Speichers vorausgeht. Wir können hierbei, wie dies Wolff (1977) auch gesehen hat, von einem Programmierungsprozeß sprechen, dessen Resultat

ein die Entnahme der aufgabenrelevanten Information steuerndes Programm ist.

Die Fourieranalyse beginnt mit der dimensional unreduzierten Verteilungsfunktion, die der Gesamtrepräsentation entspricht, und bestimmt die Periode der höchsten Invarianz, indem die Verteilungsfunktion um den Anteil desjenigen variablen Merkmals reduziert wird, dessen Periode gleich der der Gesamtfunktion ist, d.h. es wird eine Binnengliederung der Gesamtrepräsentation gesucht, die sukzessive Invarianzen gleicher Ordnung enthält. Die Reduktion der Verteilungsanteile wird solange fortgesetzt, bis das Niveau erreicht ist, das eine Invarianzperiode zeigt, die der der aufgabenrelevanten Information entspricht, d.h. mit der Intention des Hörers über die Korngröße bzw. Periode der zu beachtenden Ausschnitte übereinstimmt. Dieses dann erreichte Repräsentationsniveau kann aufgrund der Konstanz der Ordnung der dort vorhandenen Invarianzen konstant fokussiert werden und ist damit einer strukturellen Identifikation zugänglich. Der Vorgang kann aber auch in umgekehrter, also aufsteigender Richtung ablaufen, wenn der Hörer seine Intention verändert und größere Ausschnitte strukturell identifizieren oder lediglich die Verteilungen höherer Ordnung ohne strukturelle Identifikation verfolgen will.

Die Fokussierbarkeit eines niedrigeren Niveaus hängt nun davon ab, wie deutlich die Binnengliederung der Repräsentation ist, die durch den Prozeß der Reduktion der Merkmalsverteilungsfunktion bewirkt wird. Bei einer kontinuierlich auf- oder absteigenden tonalen Linie konstanter Intervalle, die auf dem

ersten und dritten Ton akzentuiert ist, schafft die Periode der Akzentvariation (relative Lautstärkenvariation) eine deutliche Binnengliederung. In diesem Fall ist die Veränderung des Merkmals Tonhöhe linear, die des Merkmals Akzentvariation aber zeigt zwei relative Maxima, die mit der Position des ersten und dritten Tons zusammenfallen (vgl. Abb. 16 linker Teil). Die Superposition beider Funktionen zeigt in Abb. 16 nur die beiden relativen Maxima der Akzentvariation, so daß die Binnenglieder klar voneinander differenzierbar und damit leicht fokussierbar sind. Eine Reduktion der Superpositionsfunktion in ihre Einzelfunktionen ist also in diesem Fall nicht notwendig, wenn eine Fokussierung des ersten oder zweiten Tonpaares intendiert ist, da die Binnengliederung bereits aufgrund des Verlaufs der Superpositionskurve deutlich wird.

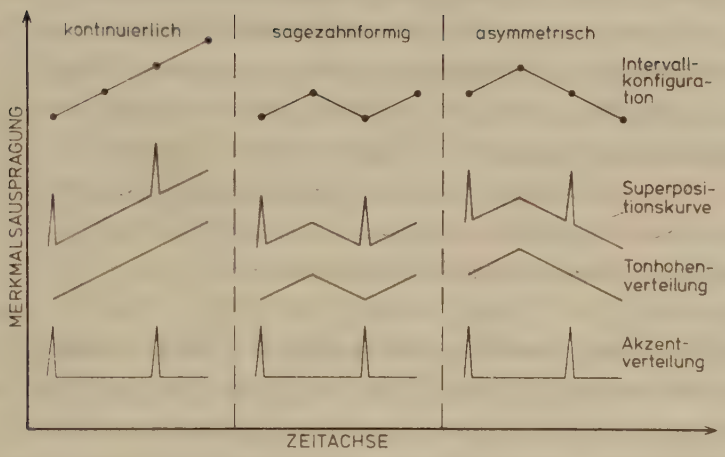


Abb. 16: Schematische Darstellung der Superposition zweier Merkmalsverteilungen in Abhängigkeit von der Intervallkonfiguration einer Tonfolge

Im Fall der mittleren Intervallkonfiguration in Abb. 16 zeigt die Superpositionskurve keine deutliche Binnengliederung, denn es liegen mehrere Maxima vor, die eine Binnengliederung in einzelne Töne nahelegen würde, wenn eine Invarianzordnung (nämlich die der Tonpaare gleicher Tonhöhe und Intervallrichtung; Invarianztyp der Wiederholung) übersprungen werden könnte. Da dann, wenn das Überspringen einer Invarianzordnung theoretisch möglich wäre, einige strukturelle Niveaus nicht fokussierbar wären, muß man annehmen, daß das kognitive System die Invarianzordnungen der Reihe nach überprüft. In diesem Fall muß also die Superpositionskurve

in ihre Komponenten reduziert werden. Dabei zeigt sich dann, daß die Minima der Tonhöhenkurve mit den Maxima der Akzentkurve zusammenfallen und an den anderen Positionen jeweils nur ein relativer Extremwert vorliegt. Die Binnengliederung aufgrund der Einzelkomponenten ist also kompatibel und zeigt Grenzmarkierungen zwischen den Paaren von Tönen. Die Kompatibilität entsteht dadurch, daß die Paare bezüglich der Intervallrichtung und der Akzentvariation invariant sind und diese beiden Invarianzen dieselbe Periode haben.

Die rechte Intervallkonfiguration besitzt eine Superpositionskurve, die ebenfalls wie beim letzten Beispiel keine deutliche Binnengliederung erkennen läßt. Die Fourieranalyse liefert nun zwei Komponenten, deren Verteilung von Extremwerten unterschiedliche Binnengliederungen nahelegen: Die Tonhöhenverteilung zeigt ein Maximum beim zweiten Ton und keine Periodik über alle vier Töne, die eine Invarianz andeutet; die Akzentverteilung zeigt wieder dieselbe Periodik und Invarianz, wie auch in den beiden anderen Beispielen. Die Tonhöhenverteilung würde eine mit der Akzentverteilung inkompatible Binnengliederung nahelegen. Da die Binnenglieder aber bezüglich eines Merkmals invariant sein müssen, kann die Binnengliederung in diesem Fall auch nur die von Paaren sein. Sie ist aber, da nicht zwei Extremwerte zusammenfallen bzw. nicht zwei Invarianzen gleicher Periode wie im letzten Fall vorliegen, undeutlicher ausgeprägt.

Wir haben unsere Vorstellung vom Abstimmungsprozeß gleich an den Beispielen entwickelt, die wir im folgenden Experiment verwenden werden, das der Überprüfung der hieraus ableitbaren Hypothesen

dienen soll. Aus unseren Aussagen über den Abstimmungsprozeß ergibt sich die folgende Hypothese über die Beziehung zwischen der Struktur einer Intervallkonfiguration und dem Zeitbedarf für die Fokussierung, die durch die Ergebnisse des Abstimmungsprozesses gesteuert wird: Wenn keine Fourieranalyse einer bereits fokussierten Invarianzperiode notwendig ist, weil die Binnengliederung in Perioden niedrigerer Invarianzordnung bereits auf der Grundlage der Superpositionsfunktion deutlich wird, sollte die Fokussierung eines aufgabenrelevanten Binnenglieds schneller möglich sein, als bei Durchführung einer Fourieranalyse. Ist jedoch eine Fourieranalyse notwendig, weil die Binnengliederung aufgrund der Superpositionsfunktion nicht deutlich wird, dann sollte die Fokussierung in den Fällen am schnellsten sein, in denen die Anzahl invarianter Merkmale zwischen Binnengliedern am größten ist, d.h. je größer die Anzahl der kompatiblen Binnengliederungen aufgrund einzelner Komponentenfunktionen ist, desto deutlicher sind die Grenzen zwischen den Binnengliedern. Einschränkend muß aber hinzugefügt werden, daß bei zu starker Binnengliederung auch ein Verlust von Objektpermanenz eintreten könnte, der dazu führen würde, daß die Gesamtrepräsentation in zwei nur als zeitlich sukzessiv repräsentierte Teile aufgespalten würde.

Daraus folgt 1., daß die kontinuierlich auf- oder absteigende Intervallkonfiguration eine schnellere Fokussierung des ersten oder zweiten Tonpaares erlauben sollte, als die sägezahnförmige Intervallkonfiguration (mittlere Konfiguration in Abb. 16), weil bei der kontinuierlich auf- oder absteigenden Konfiguration keine Fourieranalyse notwendig ist, während dies bei der sägezahnför-

migen aber der Fall ist, und 2. folgt daraus, daß die asymmetrische Intervallkonfiguration (rechte Konfiguration in Abb. 16) wegen undeutlicherer Binnengliederung als die sägezahnförmige Konfiguration die relativ langsamste Fokussierung eines der beiden Tonpaare bewirkt; und 3. schließlich folgt hieraus auch noch, daß eine Variation physischer Merkmale, die die Binnengliederung deutlicher werden läßt, die Geschwindigkeit der Fokussierung erhöhen sollte. In Grenzen dürfte also eine relativ größere Pause zwischen den Tonpaaren als innerhalb sowie ein relativ größerer Tonhöhenunterschied zwischen Paaren eine die Fokussierung erleichternde Funktion besitzen, es sei denn, die Trennung der Paare würde so groß, daß die Permanenz der Gesamttonfolge als Wahrnehmungsgegenstand verlorenginge.

Identifikation struktureller Invarianzen

Nach erfolgter Abstimmung der Periodenlänge einer Invarianzordnung können die resultierenden Binnenglieder fokussiert werden, d.h. es kann ein Repräsentationsniveauwechsel stattfinden, und diese neue Repräsentationsebene kann nun konstant fokussiert werden. Beim konstanten Fokussieren wird aber nicht automatisch eine Strukturidentifikation geleistet: Die meisten, insbesondere musikalisch wenig geschulten Hörer verfolgen lediglich die Variationen der physischen Merkmale eines Musikstücks, ohne gleichzeitig die Variationsstrukturen auf ein Kategoriensystem tonalfiguraler Muster oder gar musikalischer Formprinzipien abzubilden.

Ein Identifikationsprozeß, der die analoge Form der Repräsentation auf Kategorien abbildet, also der analog-zeitlichen Repräsentation

eine in kategorialer, nicht-zeitlicher Form zuordnet ("conceptual mapping", Laske, 198o), setzt beim Hörer das Vorhandensein differenzierter kognitiver Strukturen voraus, die den Abstimmungsprozeß den Intentionen des Hörers entsprechend steuern können und selbst wieder in Abhängigkeit von dessen Ergebnissen, also von der ermittelten Binnengliederung, ausgewählt und in Abhängigkeit von den Ergebnissen des Identifikationsprozesses modifiziert werden.

Kunst (1978) bezeichnet ein solches System kognitiver Strukturen als "prozedurale Syntax". Die Verwendung des Begriffs "Syntax" verweist darauf, daß es sich um ein System intern regelhaft organisierter Identifikationskategorien handelt. Da dieses System kognitiver Strukturen auf eine Hierarchie von Invarianzen unterschiedlicher Ordnung angewandt werden soll, ist zu fordern, daß seine interne Organisation ebenfalls hierarchisch und seine Identitätskategorien Invarianzkategorien sind. Minsky's (1975) Modell eines hierarchisch verschachtelten Systems von "frames", das ein Modell der visuellen Identifikation von Szenen und Objekten unter Rückgriff auf das Wissen des Wahrnehmenden über die Strukturen seiner Umwelt ist und eine etwas detailliertere Variante eines Schemamodells darstellt (vgl. Neisser, 1967, 1976), gibt uns einige Anhaltspunkte für eine Rahmenkonzeption der zu modellierenden Identifikationsprozesse, die mit einigen Zusatzannahmen, die sich aus unserer Konzeption der analogen Repräsentation von Tonfolgen und Musik und dem Modell der Abstimmung ergeben, zu einem Modell der strukturellen Identifikation führen.

Minsky's "frames" sind Repräsentationen des Wissens über die Strukturen der Umwelt des Wahrnehmenden mit Leerstellen oder terminalen Elementen, die von Daten aus der Umwelt besetzt werden oder selbst wieder "subframes" darstellen. Subframes repräsentieren Ausschnitte aus wahrgenommenen Situationen und können entweder Leerstellen für Daten oder weitere Subframes entsprechender Korngröße enthalten. Wir nehmen an, daß die vertikale Verschachtelung von Subframes der Hierarchie von Invarianzordnungen entspricht, die sich aufgrund der Hörerfahrungen als stabile Ordnungen musikalischer Strukturen und Teilstrukturen entwickelt haben (zur Entwicklung und Modifikation kognitiver Schemata vgl. Neisser (1976) und Piaget (1952)).

Der Identifikationsprozeß besteht nun darin, daß der fokussierte Ausschnitt eines Repräsentationsniveaus auf seine invarianten Strukturen hin untersucht wird, indem ein Subframe gewählt wird, dessen Korngröße des zu repräsentierenden Ausschnitts in Abhängigkeit von den Ergebnissen des Abstimmungsprozesses gewählt wird.

Wir entwickeln unsere Überlegungen nun wieder anhand der in Abb. 16 dargestellten Beispiele dreier Intervallkonfigurationen von vier Tönen. Die Analyse der Invarianzstruktur beginnt mit Invarianzen nullter Ordnung (alle physischen Merkmale sind konstant) und schreitet dann in aufsteigender Reihe der Invarianzordnungen fort zu der für die analysierte Korngröße höchst möglichen Invarianzordnung, sofern nicht vorher bereits eine Identifikation durch Übereinstimmung mit einer der zur Verfügung stehenden Sub-

frames erreicht wurde. Die interne Ordnung eines Frames oder Subframes bestimmter Korngröße entspricht also genau der Ordnung der Invarianzen, die in Strukturen der gerade zu analysierenden Korngröße potentiell auftreten können; zur vertikalen Schichtung der Frames und Subframes kommt also eine korngrößenspezifische horizontale Ordnung nach Invarianzordnungen.

Im Beispiel der kontinuierlich auf- oder absteigenden Tonfolge konstanter Intervalle (linke Intervallkonfiguration in Abb. 16) liegt eine Invarianz zweiter Ordnung vor, denn alleine die Merkmale Tonhöhe und Akzentverteilung variieren. Wenn die Struktur eines Subframes mit der analogen Repräsentation dieser Invarianz übereinstimmt, ist die Tonfolge als Konfiguration identifiziert. Die sägezahnförmige Tonfolge (mittlere Intervallkonfiguration in Abb. 16) besitzt keine Invarianz zweiter Ordnung, denn Tonhöhe, Akzentverteilung und Intervallrichtungen variieren; aber sie besitzt eine Invarianz dritter Ordnung, nämlich die der Wiederholung der Intervallrichtungen zwischen erstem und zweitem Tonpaar. Die asymmetrische Tonfolge (rechte Intervallkonfiguration in Abb. 16) besitzt auch keine Invarianz dritter Ordnung, denn die Beziehung zwischen den Tonpaaren zeigt eine Änderung der Intervallrichtung.

Aus dieser Darstellung folgt die Hypothese, daß der Identifikationsprozeß für die Gesamttonfolgen der drei Beispiele um so längere Zeit in Anspruch nehmen sollte, je höher die niedrigste zu identifizierende Invarianz einer Tonfolge ist, die die aufgabenrelevante Information enthält, d.h. die kontinuierlich auf- oder absteigende Tonfolge sollte am schnellsten, die asymmetrische am langsamsten identifiziert werden. Diese Hypothese steht im Widerspruch zu der Ortmann's (1933), wonach die Identifikationsschwierig-

keit von der Anzahl der Intervallrichtungsänderungen in einer Tonfolge abhängen sollte. Experiment 2 wird zwischen diesen beiden Hypothesen entscheiden können.

Während die Fokussierung eines Teils der Tonfolgen durch die Vergrößerung der Deutlichkeit in ihrer Binnengliederung bei relativ großen Pausen und/oder relativ großen Tonhöhenunterschieden erleichtert werden sollte, sollte die Identifikation der Konfiguration aller vier Töne hierdurch erschwert bzw. die Dauer des Identifikationsprozesses verlangsamt werden, denn hierdurch vergrößert sich die Variabilität der Merkmale über die ganze Tonfolge, so daß die vorhandenen Invarianzen um eine bzw. zwei Ordnungen erhöht werden, was eine Verlangsamung des Identifikationsprozesses bewirken sollte.

Fassen wir unsere Hypothesen über die Verarbeitungsgeschwindigkeiten bei der Identifikation der drei Intervallkonfigurationen bzw. des ersten oder zweiten Tonpaares zusammen, dann müssen wir in Experiment 2 die folgenden Hypothesen überprüfen:

1. Unabhängig davon, ob die Konfiguration aller vier Töne oder nur die eines der beiden Tonpaare identifiziert werden muß, sollte die Verarbeitungsdauer bei der kontinuierlich auf- oder absteigenden Tonfolge am geringsten und die bei der asymmetrischen am längsten sein.

2. Ist die Gesamttonfolge bezüglich ihrer Konfiguration zu identifizieren, dann sollte diese Identifikation weniger Zeit in Anspruch nehmen als eine Identifikation eines Tonpaares, weil unter der Voraussetzung, daß die Gesamtrepräsentation das am

Ende der Tonfolge fokussierte Niveau darstellt, im Fall der Identifikation des Tonpaares eine Abstimmung auf die Glieder der Gesamtrepräsentation und ein Repräsentationsebenenwechsel notwendig wird, der bei Identifikation der Gesamttonfolge nicht in Frage kommt.

3. In engen Grenzen sollte eine Erhöhung der Pausenlänge und/oder der Größe des Intervalls zwischen Tonpaaren im Vergleich zur Pausenlänge und Intervallgröße innerhalb der Tonpaare die Identifikation der Gesamtkonfiguration aller vier Töne verlangsamen, die der Paare aber beschleunigen.

6.2 METHODE

Versuchsplan

Wir führten ein Experiment durch, in dem als abhängige Variable die Reaktionszeit (RT) für same/different-Vergleiche zwischen Intervallkonfigurationen erhoben wurde. Zur Überprüfung der Hypothese über den Abstimmungsprozeß benötigen wir eine Bedingung, in der Ausschnitte einer Tonfolge fokussiert und identifiziert werden müssen, und zur Überprüfung der Hypothese über den Identifikationsprozeß eine Bedingung, in der die Gesamttonfolge identifiziert werden muß. Wir bezeichnen diese Bedingungen als "Gesamtaufgabe" (Identifikation der Konfiguration aller vier Töne) und als "Teilaufgabe" (Identifikation der Richtung des 1. Intervalls: Teilaufgabe 1; Identifikation der Richtung des 2. Intervalls: Teilaufgabe 2). Diese drei Aufgabenbedingungen wurden kombiniert mit den drei Intervallkonfigurationen "kontinuierlich auf- oder absteigend"

(keine Intervallrichtungswechsel), "sägezahnförmige Konfiguration" (zwei Intervallrichtungswechsel) in der in Abb. 16 dargestellten Form und deren Spiegelbild (Umkehrung der Intervallrichtungen) sowie mit der Konfiguration "asymmetrischer Verlauf" (ein Intervallrichtungswechsel) und deren Spiegelbild. Zusätzlich wurde das SOA zwischen Tonpaaren variiert (35o msec vs. 7oo msec; 35o msec SOA innerhalb der Tonpaare) und die Intervallgröße Δf zwischen Tonpaaren (kleine Terz $\hat{=}$ 3 Halbtöne vs. Quinte $\hat{=}$ 7 Halbtöne; innerhalb Tonpaaren jeweils kleine Terz).

Zusammen mit den same- und different-Bedingungen ergibt dies einen fünffaktoriellen Versuchsplan mit abhängigen Bedingungen: 3 (Intervallkonfigurationen) X 3 (Aufgaben) X 2 (SOAs) X 2 (Δ f) X 2 (same/different). Jede dieser 72 Bedingungen wurde sechsmal durchgeführt.

Das same/different-Paradigma ist deshalb hier die Methode der Wahl, weil in einer absoluten Identifikationsaufgabe, die hier eigentlich nahegelegt wird, bei Gesamtaufgabe drei Reaktionsalternativen (drei Intervallkonfigurationen), bei Teilaufgabe dagegen nur zwei Reaktionsalternativen vorliegen würden (auf- oder absteigendes Intervall eines Tonpaares). Bei same/different-Vergleich bleibt jedoch die Reaktionsalternativenzahl für alle Bedingungen konstant.

Reizdarbietung und Aufgabenstellung

Die Tonfolgen wurden wie in Experiment 1 von einem prozeßrechnergesteuerten Tongenerator erzeugt und über Kopfhörer den Vpn dargeboten. Im Unterschied zu Experiment 1 verwandten wir vier Töne von

2oo msec Dauer, wobei der erste und dritte akzentuiert wurden (5 db (A) größere Lautstärke). Aus dem Frequenzbereich zwischen 44o und 172o Hz wurde der jeweils erste Ton einer Tonfolge zufällig gewählt und die weiteren den Versuchsbedingungen entsprechend konstruiert.

Die Vergleichsaufgabe wurde über ein kleines Pult gestellt, das durch Lämpchen signalisierte, welche Aufgabe (Gesamtaufgabe, Teilaufgabe 1 oder Teilaufgabe 2) zu lösen ist und mit welcher Intervallkonfiguration die zuvor gehörte Tonfolge zu vergleichen ist (bei Gesamtaufgabe die drei Intervallkonfigurationen aller vier Töne, bei Teilaufgabe auf- oder absteigendes Intervall des ersten oder zweiten Tonpaares).

Das Pult (21 cm Breite X 16 cm Höhe der um 3o° geneigten Oberfläche) enthielt in der oberen Reihe drei schnelle Leuchtdioden: links (rot) für Gesamtaufgabe und Vergleich der gehörten Tonfolge mit der Intervallkonfiguration "kontinuierlich auf- oder absteigend", in der Mitte (grün) für Gesamtaufgabe und Vergleich der Tonfolge mit der Konfiguration "sägezahnförmig" und rechts (gelb) für den Vergleich mit der Konfiguration "asymmetrisch" bei Gesamtaufgabe. In der unteren Reihe befanden sich zwei Paare von Leuchtdioden für die Teilaufgaben: links für Teilaufgabe 1 und rechts für Teilaufgabe 2; bei Vergleich mit dem aufsteigenden Intervall leuchtete das jeweils links stehende rote Lämpchen, bei Vergleich mit dem absteigenden Intervall das jeweils rechts stehende grüne Lämpchen. Die Leuchtdioden eines Paares besaßen einen Abstand von 4.5 cm, der Abstand zwischen Paaren betrug 9 cm. Über jedem Lämpchen des Pults war die jeweilige Intervallkonfiguration in ihren beiden

zur Zeitachse spiegelbildlichen Formen für die Gesamtaufgabe in der in Abb. 16 gezeigten Form aufgezeichnet, für die Teilaufgabe die jeweilige Intervallrichtung.

Versuchsablauf

Der Versuch wurde in drei Sitzungen von jeweils ca. 45 Min. Dauer durchgeführt. Die erste Versuchssitzung diente alleine dem Training der Vpn. Zuerst wurde die Gesamtaufgabe geübt, und zwar wurden zehn same-Vergleiche und dann jeweils fünf different-Vergleiche mit immer derselben Paarung von Intervallkonfigurationen durchgeführt, bis alle Paarungen vorgekommen waren. Anschließend wurde die Paarung der Intervallkonfigurationen zufällig gewählt. Danach wurde in ebensolcher Systematik die Teilaufgabe geübt. Nach einer zehnminütigen Pause in der Mitte der Versuchssitzung wurden Teil- und Gesamtaufgaben und alle anderen experimentellen Bedingungen in Zufallsfolge dargeboten.

Während des Trainings erhielten die Vpn Rückmeldung, indem neben den Reaktionstasten befindliche Lämpchen die Richtigkeit oder Falschheit der Reaktion signalisierten (rotes Lämpchen links für falsch, grünes Lämpchen rechts für richtig).

Der Versuch wurde immer mit zwei Vpn gleichzeitig durchgeführt, die sich gegenseitig aber nicht beobachten konnten. Eine Vp hatte für same-Vergleiche die rechte und different-Vergleiche die linke Taste zu drücken, bei der anderen Vp war die Anordnung umgekehrt.

In der zweiten und dritten Versuchssitzung, in der die Daten der experimentellen Bedingungen erhoben wurden, erhielten die Vpn keine Rückmeldung mehr. Sämtliche Versuchsbedingungen wurden in Zufallsfolge dargeboten.

Die Instruktion für das Training enthielt bildlich dargestellte Beispiele der Intervallkonfigurationen und ihrer Paarungen bei same- und different-Vergleichen. In der ersten Hälfte des Trainings wurde betont, daß es alleine auf die Richtigkeit des Vergleichsurteils ankomme, in der zweiten Hälfte wurden die Vpn dann instruiert, so schnell wie möglich, aber ohne Fehler zu reagieren. Vpn, die mehr als 10 % Fehler während der zweiten Hälfte des Trainings machten, wurden von den weiteren Versuchssitzungen ausgeschlossen (dies kam nur in vier Fällen vor).

In der zweiten und dritten Versuchssitzung wurde in der Instruktion die Aufgabe noch einmal geschildert und betont, daß es sowohl auf Richtigkeit der Vergleichsentscheidungen als auch auf möglichst schnelle Reaktionen ankomme.

Die Darbietung einer Tonfolge wurde von einem 200 msec dauernden 2500 Hz Warnton eingeleitet, dem eine 1.5 sec lange Pause folgte. Dann folgten die vier Töne der Tonfolge, und unmittelbar mit Ausschalten des letzten Tons leuchtete eine der Leuchtdioden auf dem Pult auf, das die Aufgabenart und die Vergleichskontur definierte. Die RT-Messung begann unmittelbar nach Aufleuchten des Lämpchens und wurde durch Betätigung einer Reaktionstaste der Vpn gestoppt. Nachdem beide Vpn reagiert hatten, trat eine Pause von 5 sec ein, deren Ende durch den nächsten Warnton angezeigt wurde.

Same- und different-Bedingungen, die drei Aufgabenbedingungen sowie die beiden spiegelbildlichen Formen jeder Intervallkonfiguration kamen jeweils gleich häufig vor. Bei different-Bedingungen wurden bei Gesamtaufgabe die Kombinationen von Intervallkonfigurationen und bei Teilaufgabe die Kombinationen der Intervallrichtungen gleich häufig dargeboten.

Da der Prozeßrechner die Reaktionen der beiden Vpn nicht gleichzeitig verarbeiten kann, tritt bei Reaktionen der Vpn, die sich im Bereich von 0 bis 5 msec unterscheiden, ein Meßfehler in dieser Größe auf. Wir gingen deshalb davon aus, daß bei RT-Differenzen zwischen Vpn von weniger als 6 msec diese auf die serielle Abarbeitung ihrer Reaktionen zurückzuführen ist. Deshalb mittelten wir die RT-Werte beider Vpn in solchen Fällen und ordneten ihnen jeweils denselben Wert zu.

Versuchspersonen

Als Vpn nahmen am Experiment (zweite und dritte Sitzung) 27 Studenten verschiedener Fachrichtungen teil. Alle Vpn spielten seit mehr als drei Jahren ein Musikinstrument und hatten bereits am in Teil II beschriebenen Syntaxtraining teilgenommen. Alle Vpn sind also als hochtrainierte, musikalisch kompetente Hörer einzustufen. Die Vpn wurden mit 8 DM pro Versuchssitzung bezahlt.

6.3 BEFUNDE UND DISKUSSION

Auswertungsverfahren

Wir rechneten eine fünffaktorielle Varianzanalyse für Meßwiederholungen auf allen fünf Faktoren (s.Anh. 3/2). Aus denselben Gründen wie bei der Auswertung der Daten von Experiment 1 betrachten wir einen F-Wert als signifikant, wenn die mit ihm verbundene

Irrtumswahrscheinlichkeit 2 % bei regulären Freiheitsgraden und 1o % bei konservativ reduzierten Freiheitsgraden nach Greenhouse & Geisser (1959) nicht überschreitet, und als sehr signifikant betrachten wir solche Effekte, die sowohl bei regulären Freiheitsgraden als auch bei konservativ reduzierten Freiheitsgraden eine Irrtumswahrscheinlichkeit von 1 % nicht überschreiten.

Die mittlere Fehlerzahl über alle Bedingungen und Vpn betrug 4.2 %.

Haupteffekt der Intervallkonfiguration

Hypothese 1 macht die Vorhersage, daß bei kontinuierlicher Intervallkonfiguration eine schnellere Identifikation der Gesamtkontur oder eines der Intervalle möglich sein sollte als bei sägezahnförmiger Intervallkonfiguration und daß dies bei sägezahnförmiger schneller als bei asymmetrischer der Fall sein sollte.

Abb. 17 zeigt die Haupteffektmittelwerte der RT für den Intervallkonfigurationseffekt in Abhängigkeit von der Anzahl der Richtungswechsel der Intervalle einer Konfiguration.

Dieser Haupteffekt ist sehr signifikant ($F(2,52) = 7o.61$, $p < o.ol$, bei df = 1,26 ebenfalls $p < o.ol$) und alle möglichen Differenzen zwischen den drei Mittelwerten sind auf dem 1 %-Niveau signifikant (Scheffé-Test). Wie Abb. 17 zeigt, wird unsere Hypothese deutlich

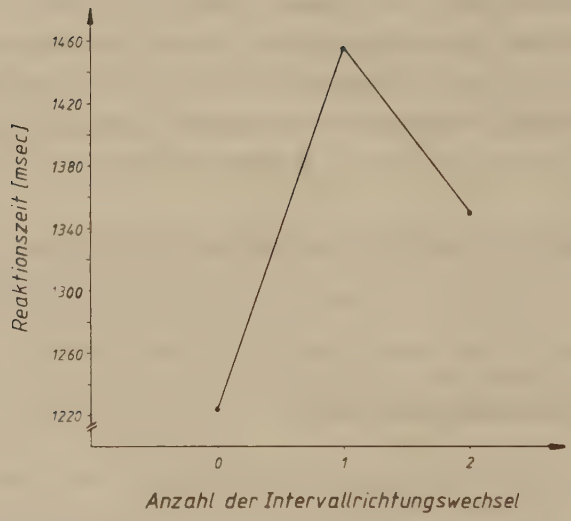


Abb. 17: RT-Mittelwerte für die Intervallkonfigurationen in Abhängigkeit von der Anzahl der Intervallrichtungswechsel

bestätigt: Die RT für die kontinuierliche Intervallkonfiguration (0 Intervallrichtungswechsel) ist kürzer als die für sägezahnförmige Konfiguration (2 Intervallrichtungswechsel) und diese wiederum kürzer als die für die asymmetrische Konfiguration (1 Intervallrichtungswechsel). Ortmann's (1933) Vermutung, wonach die Schwierigkeit der Identifikation der Struktur einer Tonfolge eine Funktion der Anzahl der Intervallrichtungswechsel sein sollte, kann durch unsere Daten nicht bestätigt werden. Vielmehr zeigt sich hier, daß die Identifikationsgeschwindigkeit von der Ordnung der Invarianz einer Tonfolge abhängt.

Effekt des Faktors "Gesamt-vs. Teilaufgabe"

Hypothese 2 macht die Vorhersage, daß die Gesamtaufgabe, also die Identifikation der Kontur aller vier Töne, schneller als die Teilaufgaben zu lösen sein sollte. Abb. 18 zeigt die RT-Mittelwerte für den Haupteffekt "Gesamt-vs. Teilaufgabe".

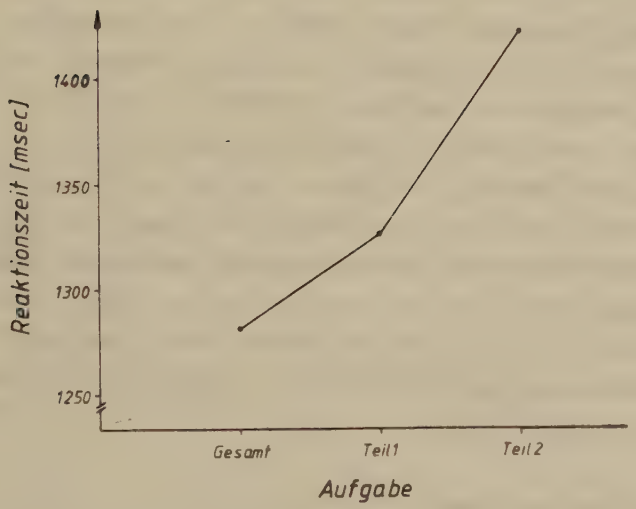


Abb. 18: RT-Mittelwerte für den Haupteffekt "Gesamt-vs. Teilaufgabe".

Dieser Haupteffekt ist signifikant ($F(2,52) = 6.84$, $p<0.01$, bei df = 1,26 $p<0.05$). Scheffé-Tests für Mittelwertdifferenzen zeigen aber, daß die RT-Differenz zwischen Gesamtaufgabe und Teilaufgabe 1 nicht signifikant ist ($p>0.10$), während die Differenz der Mittelwerte zwischen Teilaufgabe 1 und 2 ($p<0.05$), die zwischen Gesamtaufgabe und Teilaufgabe 2 ($p<0.01$) und die zwischen Gesamtaufgabe und dem Mittelwert aus beiden Teilaufgaben ($p<0.05$) signifikant sind.

Über den RT-Unterschied zwischen Teilaufgabe 1 und 2 haben wir keine Hypothese gewagt. Wenn eine serielle Positionskurve zu erwarten sein sollte, dann wäre ein "recency effect" am wahrscheinlichsten (Leshowitz & Hanzi, 1974), d.h. auf unseren Versuch übertragen, daß die Information des letzten Intervalls am leichtesten verfügbar sein sollte und deshalb auch die Identifikation des letzten Intervalls schneller als die des ersten sein sollte. Dies wäre aber nur dann zu erwarten, wenn das letzte Intervall direkt fokussiert werden könnte. Muß dagegen aber auf einem Repräsentationsniveau eine sequentielle Fokussierung durchgeführt werden, dann sollte die Identifikation des ersten Intervalls schneller als die des zweiten ablaufen können, weil hierbei ein Fokussierungsschritt gespart wird. Unsere Daten sprechen eher für diese letzte Version.

Warum der Haupteffekt keinen signifikanten Unterschied zwischen Gesamtaufgabe und Teilaufgabe 1 zeigt, geht aus den Mittelwerten der Interaktion "Intervallkonfiguration X Aufgabe" hervor. Diese Interaktion ist signifikant ($F(4,104) = 5.29$, $p<0.01$, bei df = 1,26 $p<0.05$) und zeigt (vgl. Abb. 19), daß bei der kontinuier-

lichen Intervallkonfiguration ein starker RT-Anstieg von der Gesamtaufgabe zur Teilaufgabe 1 besteht, der bei der sägezahnförmigen Konfiguration verschwindet und bei der asymmetrischen Konfiguration schließlich in eine negative Steigung übergeht, während die Unterschiede zwischen Teilaufgabe 1 und 2 beinahe konstant bleiben.

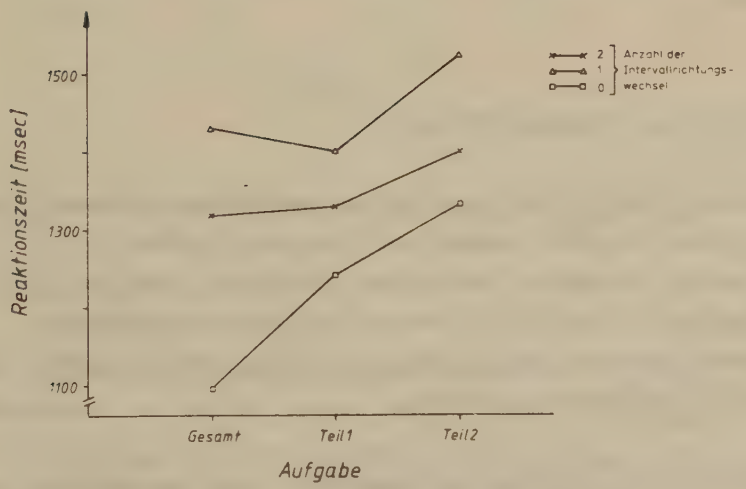


Abb. 19: RT-Mittelwerte für die Interaktion "Intervallkonfiguration X Aufgabe"

Dieser Wechselwirkungseffekt ist Ausdruck dafür, daß die Identifikationszeit mit zunehmender Invarianz der Intervallkonfigura-

tion auf dem Niveau der Gesamtrepräsentanz stärker steigt als auf dem Niveau der unmittelbar untergeordneten Repräsentationsebene. Dies ist auch kein Wunder, denn die Identifikationszeit für die Richtung eines Intervalls bei der Teilaufgabe sollte annähernd konstant bleiben, da bei dieser Korngröße keine Komplexitätsunterschiede vorliegen, die die Identifikationszeit beeinflussen könnten. Die Erhöhung der RT bei Teilaufgabe in Abhängigkeit von der Intervallkonfiguration geht also alleine auf den Abstimmungsprozeß zurück, während bei der Gesamtaufgabe die RT-Erhöhung auf den Identifikationsprozeß zurückzuführen ist.

Der Effekt der Faktoren SOA und $\triangle f$

Hypothese 3 machte die Vorhersage, daß bei Gesamtaufgabe eine Vergrößerung des SOAs und/oder von $\triangle f$ zwischen Tonpaaren relativ zum SOA und/oder $\triangle f$ innerhalb der Tonpaare zu einer Erhöhung der RT führen sollte, während bei Teilaufgabe genau das Umgekehrte eintreten sollte, sofern SOA und $\triangle f$ nicht zu einem Verlust an Objektpermanenz führen.

Einen mit dieser Hypothese konformen Effekt zeigt die sehr signifikante Interaktion "Aufgabe X SOA" ($F(2,52) = 29.72$, $p<o.ol$, bei df = 1,26 ebenfalls $p<o.ol$), die in Abb. 2o dargestellt ist.

Bei Gesamtaufgabe steigt die RT bei Verdopplung des SOAs zwischen Tonpaaren, bei Teilaufgabe 1 und 2 fällt sie um fast den gleichen Betrag. Aufgrund dieser Datenkonfiguration sollte dann kein SOA-Haupteffekt bestehen, was auch tatsächlich nicht der Fall ist

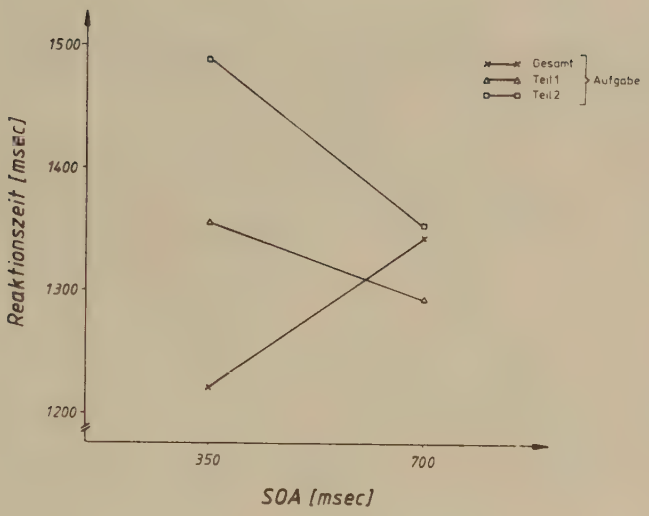


Abb. 2o: RT-Mittelwerte für die Interaktion "Aufgabe X SOA"

($F(1,26) = 4.o8$, $p \leq o.o5$).

Abb. 21 zeigt den Interaktionseffekt "Aufgabe X Δf", der sehr signifikant ist ($F(2,52) = 55.7o$, $p \leq o.ol$, bei df = 1,26 ebenfalls $p \leq o.ol$). Die Interaktion zeigt wie vorhergesagt einen Anstieg der RT bei Δf = 7 Halbtönen (Quinte) gegenüber Δf = 3 Halbtönen (kleine Terz, dies ist auch das Intervall innerhalb der Tonpaare), aber nicht den erwarteten Abfall bei Teilaufgaben, sondern vielmehr einen von Δf unabhängigen Verlauf (s. dagegen Interaktion "Aufgabe X same/different X Δf", Abb. 28, S.229).

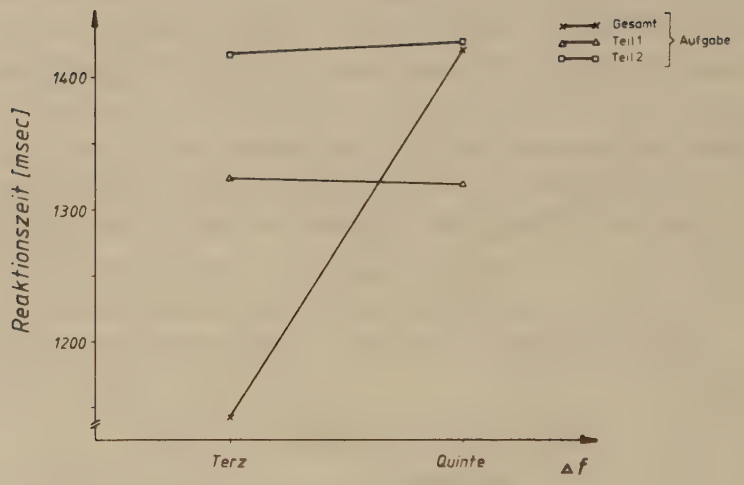


Abb. 21: RT-Mittelwerte für die Interaktion "Aufgabe X Δf"

Die sehr signifikante Interaktion "Intervallkonfiguration X Aufgabe X Δf" ($F(4,104) = 11.83$, $p \leq 0.01$, bei df = 1,26 ebenfalls $p \leq 0.01$) macht deutlich, daß der erwartete Abfall der RT bei Teilaufgabe und $\Delta f = 7$ Halbtönen gegenüber $\Delta f = 3$ Halbtönen nur bei der Intervallkonfiguration zu beobachten ist, bei der die Identifikation einer Intervallrichtung am längsten dauert, nämlich bei der asymmetrischen Intervallkonfiguration (vgl. Abb. 22).

Die sehr signifikante Interaktion "Intervallkonfiguration X Δf" ($F(2,52) = 30.50$, $p \leq 0.01$, bei df = 1,26 ebenfalls $p \leq 0.01$) bestätigt, daß nur bei der asymmetrischen Intervallkonfiguration ein die RT-verkürzender Δf-Effekt auftritt.

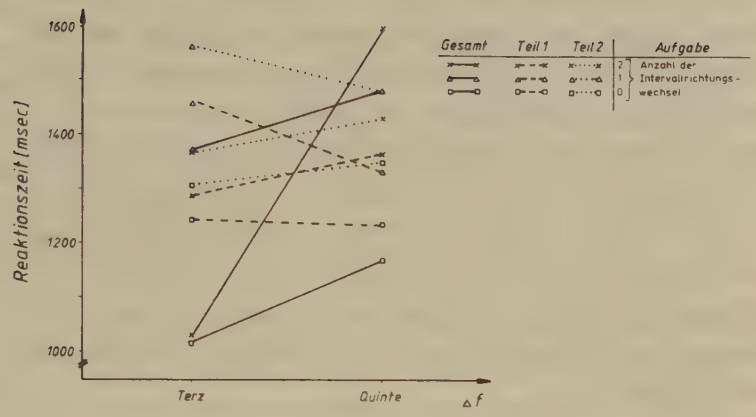


Abb. 22: RT-Mittelwerte für die Interaktion "Intervallkonfiguration X Aufgabe X Δf"

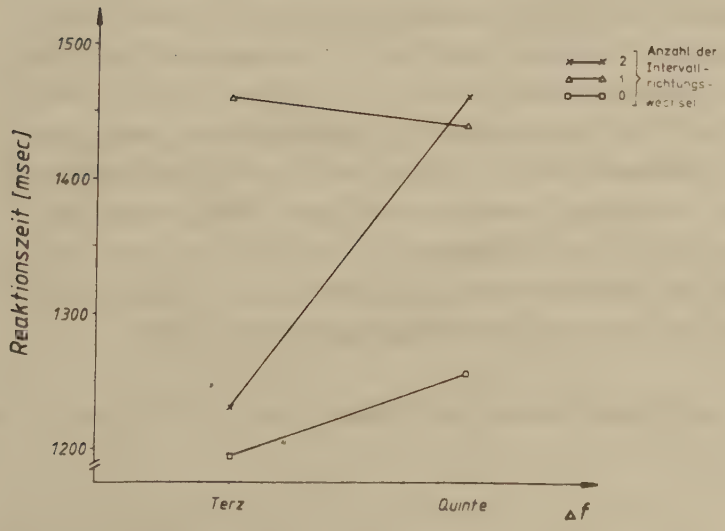


Abb. 23: RT-Mittelwerte für die Interaktion "Intervallkonfiguration X Δf"

Wir sind davon ausgegangen, daß bei Teilaufgabe durch eine geringe Erhöhung von Δf ebenso wie durch eine Erhöhung des SOAs zwischen Tonpaaren eine deutlichere Binnengliederung zwischen den zu fokussierenden Gliedern der Gesamtrepräsentation entsteht, die den Fokussierungsprozeß beschleunigen sollte. Wenn sich überhaupt eine Beschleunigung der Fokussierung nachweisen lassen sollte, dann bestimmt in der Bedingung, in der die Binnengliederung der Gesamtrepräsentation bei kleinen Δf am wenigsten deutlich ist. Dies ist bei der asymmetrischen Konfiguration der Fall und auch nur hier finden wir tatsächlich eine beschleunigte RT.

Wie bereits in den Daten von Experiment 1 sichtbar wurde, führen relativ starke Merkmalsänderungen auf der Dimension Tonhöhe zu einem Verlust der Objektpermanenz. Dies könnte auch hier die Ursache dafür gewesen sein, daß sich der Δf-Effekt bei den Konfigurationen nicht in der erwarteten Richtung zeigt, bei denen die Fokussierung bereits auf eine ausreichende Binnengliederung der Gesamtrepräsentation zurückgreifen kann, wenn Δf zwischen Tonpaaren ebenso groß ist wie innerhalb. Obgleich wir in diesem Experiment bereits eine deutlich geringere Variation von Δf als im Experiment 1 vorgenommen haben, zeigt sich hier anscheinend derselbe Effekt, wenn auch bedeutend schwächer. Daß es sich hierbei jedoch nicht um einen Effekt handeln kann, der auf Verlust der Objektpermanenz zurückgeführt werden müßte, zeigen die Interaktionen der experimentellen Variablen mit den "same/different"-Vergleichsbedingungen (S.226 - 232).

Die signifikante Interaktion "Aufgabe X SOA X Δf" ($F(2,52) = 4.33$,

$p \leq 0.02$, bei df = 1,26 $p \leq 0.10$), die in Abb. 24 dargestellt ist, macht das Befundmuster noch ein wenig komplizierter: Der nach unserer Hypothese 3 bei Teilaufgabe zu erwartende Effekt der RT-Reduktion bei Δf = 7 Halbtönen im Vergleich zu Δf = 3 Halbtönen zeigt sich hier nur bei Teilaufgabe 1 und SOA = 350 msec, während bei den anderen Teilaufgaben-SOA-Kombinationen Δf keinen Einfluß auf die RT ausübt. Die Effekte des SOAs und Δfs ergänzen sich offenbar dann nicht, wenn bereits durch ein relativ großes SOA oder ein relativ großes Δf eine Binnengliederung erzeugt wurde, die ausreichend deutlich ist, um die Fokussierung steuern zu können, d.h. die Deutlichkeit der Binnengliederung braucht nur einen Schwellenwert zu überschreiten, damit der Fokussierungsprozeß optimal ablaufen kann.

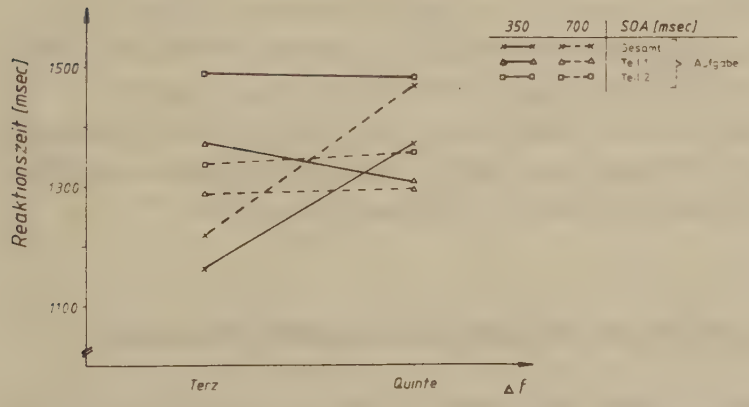


Abb. 24: RT-Mittelwerte für die Interaktion "Aufgabe X SOA X Δf"

Die signifikante Interaktion "Intervallkonfiguration X SOA" ($F(2,52) = 7.2o$, $p \leq o.ol$, bei $df = 1,26$ $p \leq o.o5$) zeigt nur für die sägezahnförmige und die asymmetrische Konfiguration eine RT-Abnahme bei SOA = 7oo msec gegenüber SOA = 35o msec, dagegen eine RT-Zunahme bei der kontinuierlichen Intervallkonfiguration (s. Abb. 25). Die Darstellung der signifikanten Interaktion "Intervallkonfiguration X Aufgabe X SOA" ($F(4,1o4) = 3.26$, $p \leq o.o2$, bei $df = 1,26$ $p \leq o.1o$) in Abb. 26 macht aber deutlich, daß dieser Effekt darauf zurückzuführen ist, daß bei Gesamtaufgabe und kontinuierlicher Intervallkonfiguration der Anstieg der RT zwischen 35o msec und 7oo msec SOA größer ist als der Abfall der RT bei Teilaufgabe, während diese Beziehung für die anderen Intervallkonfigurationen genau umgekehrt ist.

Der SOA-Effekt stellt also eine deutliche Bestätigung der Vorhersagen der Hypothese 3 dar. Der Δf-Effekt zeigt dagegen ein weniger deutliches Bild. Eine zu große Variabilität der Merkmale Tonhöhe und Intervallgröße und ein hierdurch bewirkter Verlust von Objektpermanenz sollte eigentlich in diesem Experiment ausgeschlossen sein, da ein Δf von 7 Halbtönen (Quinte) ein durchaus musikalisch häufiges Intervall ist. Der Δf-Effekt könnte aber dadurch "maskiert" worden sein, daß er bei "same"-Bedingungen einen anderen Verlauf nimmt als bei "different"-Bedingungen.

Unterschiede zwischen "same"-und "different"-Bedingungen

Die RT für "same"-Vergleiche ist sehr signifikant kürzer als die für "different"-Vergleiche ($F(1,26) = 36.35$, $p \leq o.ol$): Bei "same"-

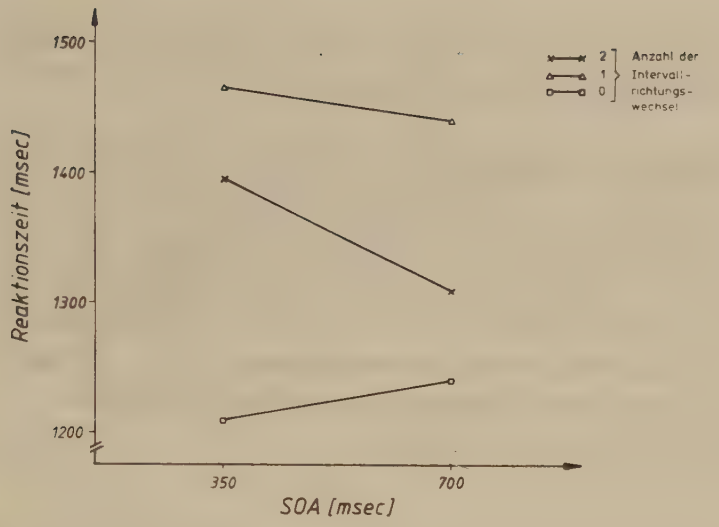


Abb. 25: RT-Mittelwerte für die Interaktion "Intervallkonfiguration X SOA"

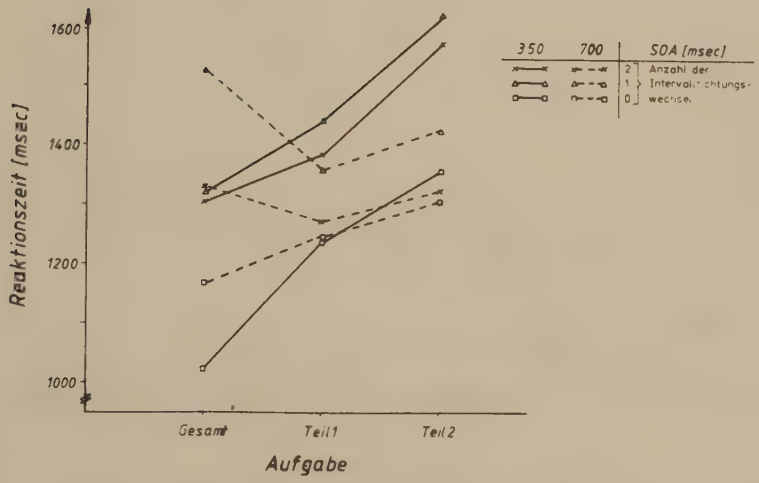


Abb. 26: RT-Mittelwerte für die Interaktion "Intervallkonfiguration X Aufgabe X SOA"

Vergleichen wurde eine mittlere RT von 1260 msec, bei "different"-Vergleichen eine von 1426 msec erreicht.

Die Interaktion "same/different X Δf" ist sehr signifikant ($F(1,26) = 8.18$, $p \leq 0.01$) und zeigt eine geringere Zunahme der RT bei "same"-Vergleichen als bei "different"-Vergleichen bei Zunahme von Δf (s. Abb. 27). Der relativ starke Anstieg der RT bei $\Delta f = 7$ Halbtönen (Quinte) im Vergleich zu $\Delta f = 3$ Halbtönen (kleine Terz), den wir auch bei Teilaufgaben fanden, was offenbar unserer Hypothese 3 widerspricht, geht also überwiegend auf die "different"-Vergleiche zurück.

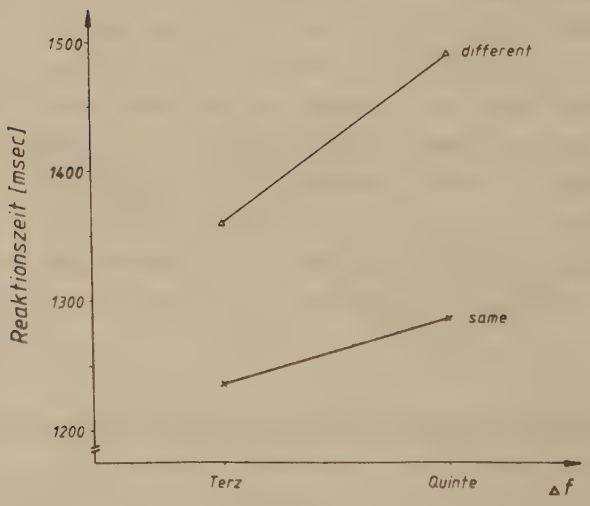


Abb. 27: RT-Mittelwerte für die Interaktion "same/different X Δf"

Ganz deutlich wird dies bei der sehr signifikanten Interaktion "Aufgabe X same/different X Δf" ($F(2,52) = 15.08$, $p \leq 0.01$, bei $df = 1,26$ ebenfalls $p \leq 0.01$), die in Abb. 28 dargestellt ist. Der in Hypothese 3 vorhergesagte Anstieg der RT bei Gesamtaufgabe und die Abnahme der RT bei Teilaufgabe bei Erhöhung von Δf zwischen Tonpaaren zeigt sich ganz deutlich bei den "same"-Vergleichen, während bei den "different"-Vergleichen eine Zunahme der RT für die Teilaufgabe zu beobachten ist.

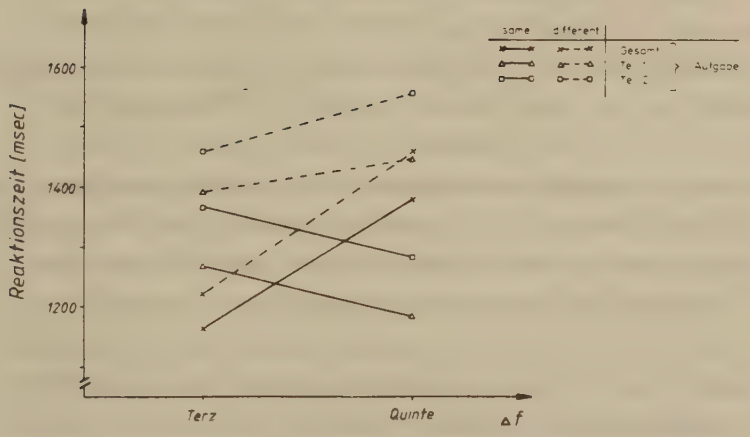


Abb. 28: RT-Mittelwerte für die Interaktion "Aufgabe X same/different X Δf"

Wenn die funktionale Basis des "different"-Vergleichs nicht, wie Bamber (1969) annimmt, in einer sequentiellen Analyse einzelner Merkmale, sondern in einer sequentiellen Analyse der V a r i a t i o n auf einzelnen Merkmalsdimensionen besteht, dann sollte 1. ein "different"-Vergleich unabhängig von allen anderen Versuchsbedingungen grundsätzlich länger dauern, und 2. könnte er auch in Abhängigkeit von der Stärke der Variabilität um so länger dauern, je größer diese ist. Da Δf = 7 Halbtöne die Variabilität des Merkmals Tonhöhe und des Merkmals Intervallgröße erhöht, könnte hierdurch die Interaktion zu erklären sein.

Hypothese 3 wird also durch die "same"-RTs voll bestätigt. Die "same"-Vergleichsbedingungen realisieren in weitaus stärkerem Maße natürliche Hörbedingungen als die "different"-Vergleichsbedingungen, denn der "same"-Vergleich ist einer absoluten Identifikation, wie sie beim normalen Musikhören stattfindet, funktional ähnlicher als der "different"-Vergleich, wenn wir Bamber's (1969) Theorie folgen, wonach "same"-Vergleiche ganzheitlich sind, also im Vergleich vollständiger Verarbeitungseinheiten bestehen, während "different"-Vergleiche auf einer elementweisen oder merkmalsweisen Überprüfung der Unterschiede beruhen.

Daß unsere Hypothesen am deutlichsten in den "same"-Vergleichsbedingungen bestätigt werden, demonstriert auch die sehr signifikante Interaktion "same/different X SOA X Δf" ($F(1,26) = 8.35$, $p \leq 0.01$). Abb. 29 zeigt, daß der SOA-Effekt bei Δf = 7 Halbtönen (Quinte) für die "same"-RTs in die erwartete, für die "different"-RTs dagegen in die entgegengesetzte Richtung geht.

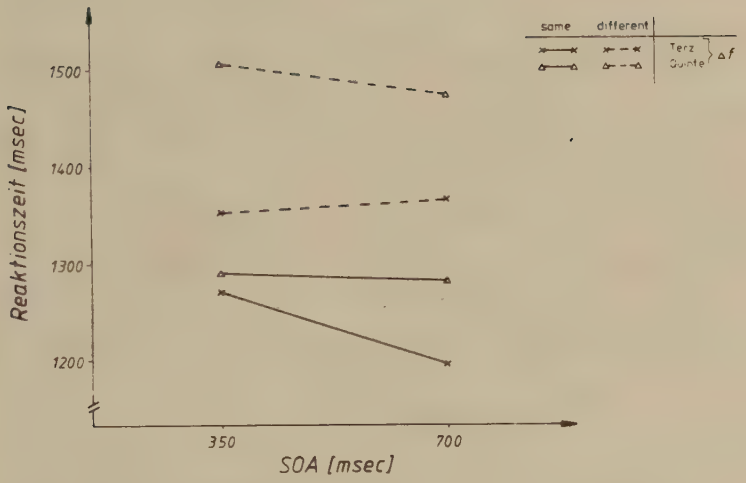


Abb. 29: RT-Mittelwerte für die Interaktion "same/different X SOA X Δf"

Ein ähnliches Bild gibt die signifikante Interaktion "Aufgabe X same/different X SOA" ($F(2,52) = 6.81$, $p \leq 0.01$, bei $df = 1,26$ $p \leq 0.05$). Abb. 3o macht deutlich, daß die in Hypothese 3 vorhergesagte längere RT bei SOA = 7oo msec im Vergleich zu SOA = 35o msec bei Gesamtaufgabe und die entsprechende RT-Abnahme bei Teilaufgabe für die "same"-Bedingungen mit den Daten übereinstimmt, während bei "different"-Bedingungen der erwartete Abfall der RT für die Teilaufgabe 1 nicht zu beobachten ist und für die Teilaufgabe 2 schwächer als in der entsprechenden "same"-Bedingung auftritt.

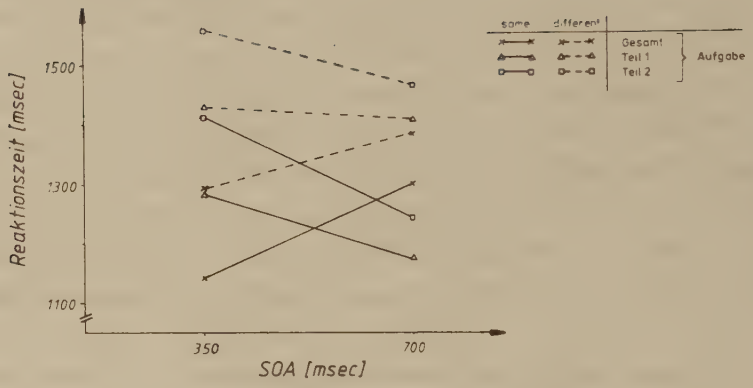


Abb. 3o: RT-Mittelwerte für die Interaktion "Aufgabe X same/ different X SOA"

6.4 ZUSAMMENFASSUNG

Aufgrund der erwartungsdiskrepanten Befunde des ersten Experiments kamen wir zu der Überzeugung, daß die erste Form der Repräsentation, die strukturell dem echoischen Speicher zuzuordnen ist, eine analoge Repräsentation sein müsse, die in ihrer vertikalen Schichtung als Hierarchie von Invarianzen unterschiedlicher Ordnung und in ihrer horizontalen (zeitlichen) Organisation als Perioden gleicher Invarianzordnungen strukturell zu beschreiben sein müßte.

Wir versuchten in Anlehnung an Gibson's "tuning"-Begriff ein funktionales Modell dieses Abstimmungsprozesses zu konzipieren, dessen Funktion darin besteht, die Fokussierung der horizontalen Periodenstruktur zu steuern, d.h. Informationen über die Invarianzordnungen und die Binnengliederung eines Repräsentationsniveaus zu gewinnen, die ein konstantes Fokussieren eines Niveaus oder einen Fokussierungswechsel (Neumann, 1980) auf ein dem gerade fokussierten Niveau über- oder untergeordneten Niveau ermöglichen. Diesen Prozeß konzipierten wir als Fourieranalyse der mehrdimensionalen Merkmalsverteilung, wie sie auf einem Niveau repräsentiert ist.

Dem Abstimmungsprozeß kann dann eine Identifikation einer der fokussierbaren Binnenglieder einer bestimmten Periode auf einem Repräsentationsniveau folgen, sofern dem Hörer geeignete kognitive Strukturen zur Verfügung stehen, die eine Abbildung der analogen Repräsentation auf strukturelle Kategorien ermöglichen. Diesen Identifikationsprozeß betrachteten wir als Invarianzanalyse, die auf ein strukturelles Wissen des Hörers zurückgreift, das in Form von Frames und hierarchisch nach Korngrößen bzw. Periodenlängen der vertikalen Schichtung der Repräsentationshierarchie organisiert ist.

Die aus diesem strukturellen und funktionalen Modell der Repräsentation und Identifikation tonaler Strukturen abgeleiteten Hypothesen über den Zeitbedarf der Identifikation von Strukturen verschiedener Korngrößen in Tonfolgen wurden im Experiment 2 überprüft. Dabei zeigte sich der erwartete Befund, daß die Identifikation unabhängig von der Korngröße der zu fokussierenden

Struktur von der Invarianzordnung des tonalen Musters abhängig ist: Je geringer die Invarianzordnung eines tonalen Musters ist, desto schneller kann es identifiziert werden. Damit konnte Ortmann's (1933) Hypothese, wonach die strukturelle Identifikation von der Anzahl der Intervallrichtungswechsel eines tonalen Musters abhängen sollte, abgelehnt werden.

Außerdem belegten die Befundmuster, daß unsere Unterscheidung zwischen Abstimmungsanalyse durch Fourieranalyse und Identifikationsprozeß durch Invarianzanalyse eine brauchbare Beschreibung der Funktionsweise der kognitiven Prozesse beim Hören von tonalen Strukturen liefert. Die entgegengesetzte Beeinflussung der Abstimmungsanalyse und des Identifikationsprozesses durch eine Erhöhung der Merkmalsvariabilität an genau der Stelle, die die Grenze zweier Binnenglieder markiert, konnte aufgrund der "same"-RT-Daten eindeutig belegt werden. Nur die Erklärung der "different"-RT-Daten bereitete einige Schwierigkeiten. Offensichtlich liegt ihnen ein anderer Funktionsmechanismus zugrunde, der im Gegensatz zur periodenweisen Identifikation bei "same"-Urteilen eine sequentielle Analyse der Variabilität einzelner Merkmale zu leisten scheint.

Das Hierarchiemodell in seiner zuletzt formulierten Version hat sich aufgrund dieser Befunde als brauchbares Modell der kognitiven Repräsentation und Identifikation tonaler Strukturen qualifiziert und eröffnet nun die Möglichkeit, die Perspektive auf Korngrößen musikalischer Strukturen zu erweitern.

7. ALLGEMEINE DISKUSSION

Struktur und Format der Repräsentation von Tonfolgen und von Musik

Am Anfang unserer Erörterung kognitionspsychologischer Probleme der Musikwahrnehmung stand die Frage, welches die spezifisch musikalischen Strukturen seien, deren Verarbeitung und Repräsentation wir in einem strukturellen und funktionalen Modell beschreiben wollten. Diese Frage zu beantworten erschien uns deshalb notwendig, weil erst eine strukturelle Beschreibung der Wahrnehmungsgegenstände die Möglichkeit eröffnet, diese systematisch zum Zwecke der experimentellen Manipulation so zu variieren, daß die Manipulationen nicht zu einem für den Gegenstandsbereich atypischen Objekt führen. Gleichzeitig kann aber eine strukturelle Beschreibung auch dazu dienen, Hypothesen über die Form der kognitiven Repräsentation zu gewinnen. Dies ist auch die Strategie, die der Entstehung propositionaler Gedächtnismodelle zugrundelag (z.B. Kintsch, 1974; Norman & Rumelhart, 1975; Rumelhart, 1975; Thorndyke, 1977) und die die Psycholinguistik so nachhaltig stimulierte, als sie begann, sich mit der Grammatiktheorie Chomsky's (1957, 1965) auseinanderzusetzen.

Doch diese Strategie verfolgten wir zunächst einmal nicht, sondern ließen uns darauf ein, das dem linearen Informationsverarbeitungsmodell meist unausgesprochen zugrundeliegende Postulat zu akzeptieren, daß die Funktion und Abfolge der kognitiven Verarbeitungsprozesse unabhängig von den spezifischen Strukturen der wahrgenommenen Reize zu beschreiben sein müßte, wenn man annimmt, daß die

Reize in einer ersten Stufe der Verarbeitung in Form elementarer Einheiten und ihrer Merkmale vorliegen, die unabhängig voneinander verarbeitet und anschließend unter Rückgriff auf Gedächtnisbestände zu einer Repräsentation konstruktiv vereinigt werden (z.B. Haber, 1967; Neisser, 1967). In diesem Fall ist dann das Format der Repräsentation überwiegend von den individuellen Gedächtnisbeständen, am allerwenigsten aber von der Struktur der Reize selbst abhängig.

Bei Annahme dieser Voraussetzungen kann man Musikwahrnehmung auch am Beispiel "unmusikalischer" Reize untersuchen. Doch die Befunde, die auf diesem theoretischen Hintergrund zur Verarbeitung und Repräsentation von lediglich nach methodischen Gesichtspunkten variierten Tonfolgen erhoben wurden, ließen nur selten eine Erklärung zu, die mit den Voraussetzungen linearer Informationsverarbeitungsmodelle zu vereinbaren waren. Insbesondere die phänomenale Ganzheitlichkeit der Wahrnehmung von Tonfolgen und die vielfach demonstrierte Abhängigkeit der Identifikation tonaler Ausschnitte und ganzer Tonfolgen von der Konfiguration ihrer Merkmale standen in krassem Gegensatz zur Annahme unabhängiger Merkmalsanalysen und linear vom Elementaren zum Komplexen fortschreitender konstruktiver Prozesse der Herstellung einer Repräsentation für eine Tonfolge.

Die Alternative zum linearen Modell sahen wir im Hierarchiemodell der kognitiven Verarbeitung, wie es von Wolff (1977) zur Erklärung der Identifikation und Repräsentation strukturierter Buchstabenfolgen entwickelt wurde. Der Kerngedanke dieses Modells besteht darin, eine Hierarchie von Repräsentationen verschiedener Korn-

größe der Reizmusterstruktur anzunehmen, die bei der Entnahme von Identitäts- und Positionsinformationen über die Reizstruktur "top-down" von der höchsten zur niedrigsten Repräsentationsebene durchlaufen wird, wobei die aufgabenrelevante Information der jeweils höchsten Repräsentanz entnommen werden kann, die diese enthält.

Die Annahme, Musik strukturell als Hierarchie von Einheiten unterschiedlicher zeitlicher Erstreckung auffassen zu können, hat spätestens seit Riemann (1889, 1914/15) Eingang in die musikwissenschaftliche Analysemethodik gefunden. Eine hierarchische Repräsentationsstruktur für Musik ist inzwischen auch mehrfach von kognitionspsychologischer Seite als Modell der Organisation musikalischer Einheiten im Gedächtnis vorgeschlagen worden (Deutsch, 1980b; Jones, 1978; Restle, 1970). Doch im Unterschied zu diesen Modellen der Gedächtnisrepräsentation musikalischer Strukturen eröffnet das Hierarchiemodell von Wolff (1977) die Möglichkeit, die Repräsentation der Wahrnehmungsinhalte und die Prozesse ihrer Analyse als hierarchisch zu begreifen: Die Repräsentationshierarchie ist nicht erst das g e s p e i c h e r t e R e s u l t a t einer aufmerksamkeitsgesteuerten Verarbeitung, sondern selbst G e g e n s t a n d dieser Verarbeitung, sie ist nicht Gedächtnis- sondern Wahrnehmungsstruktur. Davon unberührt bleibt die Hypothese einer ebenfalls hierarchischen Gedächtnisstruktur weiterhin plausibel, gewinnt sogar noch an Wahrscheinlichkeit, weil die Kodierung einer hierarchischen Gedächtnisrepräsentation auf der Basis einer bereits hierarchisch strukturierten Wahrnehmungsrepräsentation ein weitaus weniger kompliziertes Kodierungsmodell verlangt, als ihre Konstruktion aus unabhängig verarbeiteten Merkmalen einzelner Töne.

Die Befunde von Experiment 1 machten dann schließlich deutlich, daß wir es versäumt hatten, einem Aspekt der kognitiven Repräsentation Beachtung zu schenken, der auch in den meisten Informationsverarbeitungsmodellen gar nicht explizit erörtert wird. Bei der Übertragung des Modells von Wolff (1977) haben wir stillschweigend die Voraussetzung übernommen, daß die "Einheiten" der Repräsentation diskrete Entitäten sind. Die Begriffe, mit denen wir die Strukturen und ihre Repräsentation darzustellen versuchten, z.B. der Begriff der "Korngröße" und der der "Verarbeitungseinheit", sind sämtlich Konstrukte zur Beschreibung s t a t i s c h e r O b j e k t e und ihrer Merkmale bzw. deren statischer Repräsentation. Wenn aber zeitliche Variationen die Wahrnehmung angeblich statischer Einheiten systematisch beeinflussen, obwohl die Zeit nicht "Merkmal" sondern Bezugssystem der Einheiten ist, dann müssen diese Einheiten immer in Relation auf ihr Bezugssystem repräsentiert sein, das aber heißt, die Repräsentation muß im Format der zeitlichen Dynamik der Merkmale einer Tonfolge vorliegen. Es muß also über die Frage der Struktur einer Repräsentation hinaus auch noch die Frage ihres Formates beantwortet werden, und damit sind wir erneut auf das Problem verwiesen worden, daß offenbar eine Wahrnehmungstheorie nicht ohne eine Definition der die Wahrnehmungsobjekte bildenden Reizaspekte auskommt.

Der Reiz ist für eine Theorie der Wahrnehmung ebenso ein wissenschaftliches Problem wie die Mechanismen der Wahrnehmung (Gibson, 1966a; Mace, 1977). Wenn die adäquate Beschreibung der Reize erst einmal zum wissenschaftlichen Problem erhoben wird, was im Rahmen der Informationsverarbeitungsmodelle selten geschieht, könnte und

sollte man nach Mace (1974) sogar davon ausgehen, daß die Lösung dieses Problems sowohl logisch als auch strategisch jedem Versuch einer Modellierung der Verarbeitungsprozesse vorgeordnet zu sein hat. Indem man die Charakteristika der Wahrnehmungsobjekte zuerst zu beschreiben versucht, kann man die Klasse der plausiblen Mechanismen der Wahrnehmung bereits so eingrenzen, daß sie das zu leisten vermögen, was die Struktur der Wahrnehmungsobjekte an Anforderungen stellt (Mace, 1977).

Der Kernsatz, mit dem Mace (1977) Gibson's (1966a) Forschungsstrategie charakterisiert, ist ein wenig irreführend: "Ask not what's inside your head, but what your head's inside of" (Mace, 1977, S. 43). Vielmehr sollte es heißen: "Frage nicht zuallererst danach, was in Deinem Kopf vorgeht, sondern zuerst danach, in was für einer Welt sich Dein Kopf befindet, und dann stelle die Frage nach den Vorgängen in Deinem Kopf erneut."

Wichtiger als die Reihenfolge dieser Fragestellungen, die lediglich die Ökonomie einer wissenschaftlichen Forschungsstrategie betrifft, ist das Stellen und Beantworten der Frage nach den Strukturen der Wahrnehmungswelt, ganz gleich, an welcher Stelle des Forschungsprozesses dies geschieht, denn auch Mißerfolge bei der Interpretation von Daten können darauf aufmerksam machen, daß die in das Erklärungsmodell unreflektiert eingeflossenen Voraussetzungen über die Reizmusterstrukturen nicht zutreffen können, wie das Beispiel von Experiment 1 zeigt.

Die Befunde von Experiment 1, speziell die Richtung der Effekte der Variablen $\Delta f_{1,2}$ (Tonhöhenintervall zwischen 1. und 2. Ton)

und der Variablen $SOA_{1,2}$ und $SOA_{2,3}$ (SOAs zwischen 1. und 2. bzw. 2. und 3. Ton) bzw. deren Interaktionen, entsprachen nicht den Erwartungen, die wir aufgrund der Chunkingkonzeption gebildet hatten: Große zeitliche oder Merkmalsunterschiede zwischen dem 1. und den beiden folgenden Tönen führten nicht zu einer Chunkbildung zwischen den beiden letzten Tönen, die eine Entnahme der aufgabenrelevanten Intervallrichtung hätte erleichtern und dadurch die RT verringern können. Stattdessen stellte sich die kürzeste RT bei Homogenität der SOAs und bei minimaler Variabilität der Intervallgrößen und -richtungen ein.

Ein solches Befundmuster ist dann zu erwarten, wenn die Verarbeitung sensibel für die Variabilität der Merkmalsverteilung über die Zeit ist. Doch können wir diese Sensibilität nicht alleine auf die funktionalen Eigenschaften des Verarbeitungsmechanismus zurückführen, denn wie sollte es zu erklären sein, daß die Verarbeitung von der Variabilität der Merkmale über die Zeit beeinflußt wird, wenn nicht bereits diese Variabilität die Informationsbasis für den Verarbeitungsmechanismus darstellt. Wenn die Fokussierung also bei der Entnahme von Identitätsinformationen auf den strukturierten Inhalt des echoischen Speichers zugreift, dann muß das Format dieser Strukturierung bereits eines sein, das die zeitliche Variabilität der Merkmale repräsentiert. Das Format der Repräsentation muß der Variabilität der Reizmerkmale analog sein, kann also nicht die Form diskreter Einheiten annehmen.

Wenn diese Annahme über das Format der Repräsentation eine adäquate Beschreibung der Wahrnehmung darstellt, sollte sie nicht nur für

die auditive Modalität zutreffen, in der die Zeit eine so offensichtliche Rolle spielt, sondern auch für alle anderen Modalitäten. Genau dieser Auffassung ist Gibson (1966a): Die Wahrnehmung eines statischen Objekts ist lediglich ein (seltener) Grenzfall ökologischer Wahrnehmungsbedingungen (vgl. auch Neisser, 1976); Wahrnehmung besteht vielmehr in der Repräsentation der Invarianz von Objektmerkmalen und ihrer Konfiguration (strukturelle Invarianz) und in der Repräsentation der Systematik ihrer Veränderung in Raum und Zeit (transformationelle Invarianz) (Shaw & McIntyre, 1974; Shaw & Pittenger, 1978).

Dann aber kann man nicht davon sprechen, daß "Elemente" ein Objekt konstituieren, sondern es ist die sich bei Veränderung erweisende Invarianz einer Beziehungsstruktur, die ein Objekt konstituiert, an dem Elemente und Merkmale erst nach Entdecken der Invarianz, die das Objekt konstituiert, in nachträglicher Analyse zu identifizieren sind. Die Identifikation eines Ausschnitts aus einer Melodie setzt also das Entdecken einer diesen Ausschnitt konstituierenden Invarianz voraus, die sozusagen das "Objekt" darstellt, an dem ein "Element" aufgefunden werden kann. Wenn dieser Ausschnitt fokussiert wird, wird er selbst zum "Objekt", muß also selbst wieder eine Invarianz niedrigerer Ordnung enthalten. Diese "Objekt-Teil-Relationen" konstituieren damit eine Hierarchie von Invarianzen unterschiedlicher Ordnung.

Das Format der strukturellen Repräsentation einer Tonfolge haben wir als "analog" bezeichnet. Diese Charakterisierung einer Repräsentation haben wir in Anlehnung an das Konzept der "analogs" bei Neisser & Hirst (1974) vorgenommen. Neisser & Hirst (1974) stellen

den "analogs" die Repräsentation einer Tonfolge in Form von "strings" gegenüber, die auf der Basis der analogen Repräsentation kodiert werden. "Strings" sind abgegrenzte, einzeln fokussierbare Einheiten, "analogs" dagegen stellen eine nicht-diskrete Repräsentation des gesamten Reizmusters ohne Binnengliederung dar, die aufgrund der durch die Merkmalsanalyse auf präattentivem Niveau (Neisser, 1967) zur Verfügung gestellten Information so k o n s t r u i e r t werden, daß sie zur Diskrimination zwischen alternativen strukturellen Organisationen einer Menge von Tönen geeignet sind, wie dies in den Aufgaben von Neisser & Hirst (1974) erforderlich war.

In Übereinstimmung mit Neisser & Hirst (1974) gehen wir davon aus, daß vor einer diskreten Repräsentation eines Wahrnehmungsobjekts, wie sie als Ergebnis eines Identifikationsprozesses entstehen kann, in dessen Verlauf das Objekt in seiner kategorialen Zugehörigkeit zu einer Klasse von Objekten repräsentiert wird, eine Repräsentation in analoger Form vorliegen muß, auf die bei Fokussierung selektiv zugegriffen werden kann. Aber im Unterschied zu ihnen und in Übereinstimmung mit Gibson (1966a, 1977) und Neisser's späteren Ansichten (Neisser, 1976), die entschieden von Gibson geprägt wurden, sehen wir in der Herstellung einer analogen Repräsentation keinen konstruktiven Prozeß, der durch Zugriff auf Gedächtnisinhalte und unter Berücksichtigung aufgabenspezifischer Bedingungen eine Repräsentation aufbauen würde. Die Argumente für eine solche "direkte Wahrnehmung" sind den zahlreichen Darstellungen des Gibson'schen Ansatzes zu entnehmen (Mace, 1974, 1977; Neisser, 1976; Shaw & Bransford, 1977; Shaw & McIntyre, 1974; Turvey, 1975, 1977).

Der zweite Unterschied zwischen Neisser & Hirst's (1974) "analogs" und der analogen Repräsentationshierarchie in unserer Sichtweise besteht in der Vollständigkeit der Repräsentation der wahrgenommenen Tonfolge. Aufgrund der konstruktivistischen Konzeption bei Neisser & Hirst (1974) ist ein "analog" bereits eine selektive Repräsentation, die nur aufgabenrelevante Informationen über die physischen Merkmale einer Tonfolge enthält. In unserer Konzeption kann zwar auch auf die Repräsentationshierarchie selektiv zugegriffen werden, hierdurch verändert sich aber nicht der Inhalt der Repräsentation, sondern nur der Ausschnitt des gerade fokal beachteten Repräsentationsniveaus. Unsere Konzeption ist also insofern ökonomischer, als sie bei Aufgabenwechsel nicht die Neukonstruktion einer geeigneten Repräsentation verlangt, sondern lediglich einen Wechsel des Repräsentationsniveaus auf eine Ebene, der die aufgabenrelevante Information zu entnehmen ist. Während eine Konstruktion einer die aufgabenrelevante Information beinhaltenden analogen Repräsentation kaum kapazitätsfrei ablaufen kann, weil hierzu Such- und Vergleichsprozesse, eine aktive Selektion aufgabenrelevanter Merkmalsvariationen und eine kurzfristige Speicherung notwendig sind, kann ein Wechsel des Repräsentationsniveaus kapazitätsfrei konzipiert werden, so wie wir dies im Rahmen unseres Modells der Fourieranalyse getan haben. Erst eine kapazitätsfreie Konzeption des Zugriffs auf die Repräsentationen eines Wahrnehmungsobjekts erlaubt es zu erklären, warum es so leicht ist, ein Objekt einmal als Gegenstand und das andere Mal unter dem Aspekt seiner Oberflächenbeschaffenheit zu sehen oder umgekehrt.

Die konstruktivistische Konzeption einer analogen Repräsentation von Tonfolgen hat einige Schwierigkeiten zu erklären, daß ein musikalisch ungeschulter und wenig hörerfahrener Hörer überhaupt Musik hören und sich gar an ihr erfreuen kann. Besitzt auch ein solcher Hörer die Fähigkeit, zwischen verschiedenen Merkmalsvariationen zu selegieren und musikalische Informationen kurzfristig zu speichern, um sie dann zu einer Repräsentation zu integrieren? Und wenn ja, kann dann Musikhören für ihn überhaupt noch ein Genuß sein, da doch seine kognitiven Systeme auf Hochtouren laufen müssen, um Repräsentationen des Gehörten zu konstruieren? Wenn aber Wahrnehmung "direkt" ist, besteht dieses Problem überhaupt nicht. Selbst der unmusikalischste Hörer kann dann den simultanen Variationen physischer Merkmale "lauschen", ohne die geringste Anstrengung damit verbinden zu müssen. Ob ihm diese Variationen aber geordnet oder gar sinnvoll erscheinen, ist ein ganz anderes Problem, nämlich das der Identifikation oder das des Wiedererkennens bekannter Variationsmuster. Gibson unterscheidet zwischen "perceptual cognition or knowledge o f the enviroment, and symbolic cognition, or knowledge a b o u t the enviroment" (Gibson, 1966a, S. 91). Wissen v o n Musik kann jeder Hörer erlangen, Wissen ü b e r Musik setzt die Fähigkeit zur symbolischen oder kategorialen Repräsentation voraus; nur die letztere Form des Wissens ist eine kodierte Repräsentation.

Entdecken von Invarianzen und Fourieranalyse

Ein funktionales Modell, das den Zugriff und die Analyse einer analogen Repräsentationsstruktur beschreiben soll, muß dem ana-

logen Format der Information Rechnung tragen. Damit scheint die Analogie zum Digitalcomputer unbrauchbar zu werden. Die Mechanismen des Wahrnehmungssystems sind wahrscheinlich adäquater modelliert, wenn sie als Dekomposition der simultan repräsentierten Variationen auf physikalischen Merkmalsdimensionen beschrieben werden. Eine solche Form der Analyse müßte dazu in der Lage sein, kontinuierliche Informationen über beliebig viele Merkmalsdimensionen in ihre Variationscharakteristika zu zerlegen, die im einfachsten Fall in Perioden identischer Merkmalswiederholungen und in komplexeren Fällen in Perioden höherer Momente einer Veränderungsfunktion (Transformationsfunktion für Merkmalsverteilungen in der Zeit) bestehen könnte.

Wenn die Merkmalsvariationen aber für mehrere Dimensionen als integrale Variationsstruktur repräsentiert sind, dann müssen die Perioden einzelner Merkmalsvariationen voneinander isoliert werden, um ein konstantes Fokussieren eines Repräsentationsniveaus sowie einen gesteuerten Fokussierungswechsel auf ein höheres oder niedrigeres Niveau zu ermöglichen.

Ein Prozeß, der dies zu leisten vermag, könnte im Prinzip in einer Fourieranalyse der als Superposition eindimensionaler Merkmalsvariationen repräsentierten tonalen Struktur bestehen. Die Voraussetzung, die wir für ein solches Modell machen müssen, besteht in der Annahme, daß das Prinzip der Fourieranalyse, mit dem die neuralen Mechanismen des Hörens einzelner Töne beschrieben werden (Broadbent, 1975), auch auf das makroskopische Niveau der Variation zwischen Tönen angewandt werden kann. In Pribram's (1977a,b) holographischem Modell der visuellen Wahrnehmung wird angenommen,

daß nicht die einzelnen ortsfrequenzspezifischen Kanäle sondern die Interferenzmuster, die auf einer Projektionsebene als Überlagerung aller Ortsfrequenzen entstehen, auf Invarianzen analysiert werden. Diese Invarianzanalyse beschreibt Pribram (1977a,b) als Fourieranalyse. Die Fourieranalyse dient dazu, die Periodik von Invarianzen einer bestimmten Ordnung erkennbar zu machen, so daß den Intentionen des Hörers entsprechend eine Periodenlänge aufgesucht und fokussiert werden kann.

Strukturelle Identifikation durch Resonanz

Pribram (1977a,b) geht noch einen Schritt weiter und nimmt an, daß ein solches Interferenzmuster vom Gedächtnis ohne symbolische Kodierungen direkt gespeichert und durch Resonanz mit identischen oder partiell übereinstimmenden, bei aktuellen Wahrnehmungen entstehenden Interferenzmustern durch Resonanz erneut aktiviert werden kann. Ein solcher Resonanzprozeß könnte die Grundlage eines musikalischen Gedächtnismodells bilden, das ohne aktive Such- und Abrufungsprozesse auskommen könnte. Damit kann insbesondere eine Erklärung für die Leichtigkeit des Wiedererkennens größerer musikalischer Abschnitte gegeben werden. Hiermit lassen sich die bei Dowling (1978) referierten Befunde zum Wiedererkennen von transformierten Melodien mit dem Prinzip der Invarianzordnung erklären: Die Resonanz sollte um so größer sein, je geringer die Ordnung der Invarianz zwischen der Standard- und der Vergleichsmelodie ist; bei identischer Wiederholung liegt Invarianz nullter Ordnung vor, bei Transposition Invarianz erster Ordnung, bei Verschiebung Invarianz zweiter Ordnung, bei Umkehrung

der Intervallrichtungen und bei zeitlicher Umkehrung eine noch höhere Invarianzordnung und schließlich bei Kombination dieser beiden Transformationen die höchste Invarianzordnung. Die Wiedererkennungsraten bilden dieselbe Rangfolge wie die der Invarianzordnungen (vgl. unsere Zusammenstellung dieser Befunde auf S. 49). Ein solches Resonanzmodell des Gedächtnisses haben auch Bransford et al. (1977) entworfen und sich dabei explizit auf den Gibson'schen "tuning"-Begriff bezogen.

In unserer ersten Konzeption des Hierarchiemodells waren wir noch von der Voraussetzung ausgegangen, daß die elementare Identitätskategorie eine "tonale Linie" sei. Unter dieser Voraussetzung konnten wir dann den Identifikationsprozeß in den Diskriminationsaufgaben von Divenyi & Hirsh (1974, 1975) und Kerivan & Carey (1976) und den Konfigurationshaupteffekt in Experiment 1 dadurch erklären, daß wir annahmen, die Repräsentationshierarchie werde solange durchlaufen, bis auf einem Niveau ein Teilmuster als "tonale Linie" zu identifizieren sei, und für die weiteren Teilmuster werde auf eventuell noch niedrigeren Niveaus nach "tonalen Linien" gesucht.

Diese Erklärung ist aber dann wenig plausibel, wenn man mit ihr auch die absolute Identifikation beim Musikhören erfassen will. Muß ein Hörer, um zwischen verschiedenen Leitmotiven in Wagner's "Ring der Nibelungen" zu diskriminieren, das gerade gehörte Leitmotiv erst in seine Elemente konstanter Intervallrichtungen zerlegen, bevor er erkennen kann, um welches Leitmotiv es sich hierbei handelt? Oder kann der Hörer von Beethoven's "Ode an die Freude" in der 9. Sinfonie diese nur dadurch identifizieren, daß

er den ersten Takt als aufsteigende Linie und den zweiten als absteigende identifiziert? Wir glauben, daß außer vielleicht in solchen Diskriminationsaufgaben, wie die bei Divenyi & Hirsh (1974, 1975), kaum ein Identifikationsprozeß beim Musikhören sich solcher elementarer Konturmuster bedient. Zudem müßte bei der Identifikation zeitlich weitreichender Beziehungsstrukturen ein mehrfacher Repräsentationsebenenwechsel in Richtung detaillierterer Binnenstrukturen vorgenommen werden. Diese Binnenstrukturen können aber kaum noch exakt zu analysieren sein, wenn sie wegen der großen zeitlichen Erstreckung der zu analysierenden Gesamtstruktur die psychische Präsenzzeit überschreiten.

In unserer zweiten Konzeption sind wir deshalb davon ausgegangen, daß der Periode einer zu analysierenden Invarianz eines Repräsentationsniveaus entsprechend eine kognitive Struktur gewählt wird, deren hierarchisch verschachtelte Untersysteme ("subframes") in der Reihenfolge der vorfindbaren Invarianzen niedrigerer Ordnung eine strukturelle Analyse vornehmen. Das Repräsentationsniveau bleibt hierbei konstant, aber das Niveau der Invarianzanalyse durch die Subframes ändert sich in Abhängigkeit von den Invarianzordnungen, die die Binnenstruktur der Repräsentation erkennen lassen.

Z.B. das Erkennen von strukturellen Beziehungen innerhalb der musikalischen Periode kann durch einen solchen Prozeß beschrieben werden. Wenn z.B. der Vordersatz aus vier Takten innerhalb einer achttaktigen Periode gerade abgeschlossen ist und der Nachsatz beginnt, kann der Vordersatz auf dem Niveau, das eine Periode von vier Takten besitzt, mit dem sich gerade entwickelnden Nach-

satz strukturell verglichen werden, indem die Invarianzordnungen des Vordersatzes mit den sich entwickelnden Invarianzen des Nachsatzes von den niedrigen Invarianzordnungen zur höchsten mit Hilfe eines entsprechenden Framesystems analysiert werden. Wenn sich Invarianzen gleicher Ordnung und gleichen Typs innerhalb von Vorder- und Nachsatz zeigen, wird dies durch Resonanz des entsprechenden Subframes signalisiert und der Nachsatz ist hierdurch in seiner kategorialen Beziehung zum Vordersatz identifiziert (z.B. als Abwandlung der 2. Subphrase (Takt 3 und 4) des Vordersatzes in der zweiten Subphrase des Nachsatzes (Takt 7 und 8), die aus der Notwendigkeit, den Grundton zu erreichen, resultiert).

Unsere funktionalen Aussagen über die kognitiven Prozesse des Strukturerkennens in Tonfolgen sind zugegebenermaßen noch recht vage und keineswegs in all ihren Konsequenzen experimentell überprüft. Doch glauben wir hierdurch eine theoretische Basis geschaffen zu haben, die mehr als die linearen Informationsverarbeitungsmodelle der auditiven Mustererkennung dazu geeignet ist, zur Erklärung der kognitiven Prozesse der Musikwahrnehmung beizutragen.

Im folgenden zweiten Teil unserer Untersuchung werden wir uns explizit mit der Wahrnehmung und Repräsentation echter Musik beschäftigen und den strukturellen Aspekt der Repräsentation musikalischer Strukturen in den Mittelpunkt stellen. Durch experimentelle Überprüfung von Hypothesen, die aus einem formalen Modell einer musikalischen Syntax abgeleitet werden, versuchen wir ein differenzierteres Bild der kognitiven Repräsentation von Strukturen höherer Ordnung zu zeichnen, als wir dies in Teil I tun konnten.

TEIL II

WAHRNEHMUNG UND REPRÄSENTATION MUSIKALISCHER STRUKTUREN: ZUR PSYCHOLOGISCHEN BRAUCHBARKEIT EINER MUSIKALISCHEN SYNTAX

1. EINLEITUNG

Die in Form des Hierarchiemodells nun vorliegende theoretische Rahmenkonzeption zur Erklärung struktureller und funktionaler Aspekte der kognitiven Verarbeitung bei der Wahrnehmung und Identifikation von Tonfolgen soll nun auf die Wahrnehmung und Repräsentation musikalischer Strukturen angewandt werden. Diese Erweiterung der Perspektive macht es nun erneut notwendig, Klarheit über die Beschaffenheit der zum Wahrnehmungsobjekt werdenden Strukturen zu erlangen. Wir verfolgen deshalb die von Gibson (1966a) eingeschlagene Forschungsstrategie, zuerst nach den Strukturen der Wahrnehmungsobjekte zu fragen, um dann anschliessend zu prüfen, ob die Wahrnehmung und Repräsentation solcher Strukturen im Rahmen des Hierarchiemodells zu erklären ist.

Die Erfüllung methodologischer Voraussetzungen, eine solche Forschungsstrategie mit Aussicht auf Erfolg einschlagen zu können, verdanken wir der Linguistik, deren formal-analytisches Methodenrepertoire von einigen musikalisch interessierten Linguisten zur Analyse musikalischer Strukturen herangezogen worden ist. Die syntaktische Beschreibung eines musikalischen Formtyps wird deshalb den Ausgangspunkt unserer weiteren theoretischen Überlegungen bilden.

Indem wir versuchen, psychologische Fragestellungen zur Wahrnehmung von Musik zu bearbeiten, knüpfen wir an eine alte Forschungstradition an, die beinahe genau so alt ist, wie die experimentelle Pschologie selbst. Dem Anspruch, Musikhören und Musikausübung psychologisch zu erklären, konnte die Musikpsychologie bis-

lang aber kaum gerecht werden, weil sie sich durch die Einseitigkeit der thematisierten Fragestellungen in eine selbstverschuldete Isolation begeben hat, die sie den Anschluß an die theoretischen Entwicklungen in anderen Bereichen der Psychologie, insbesondere der Kognitionspsychologie, in den letzten Jahrzehnten verlieren ließ.

1.1 KOGNITIVE MUSIKPSYCHOLOGIE: EIN FORSCHUNGSPROGRAMM

Die Psychologie der Musikwahrnehmung war lange Zeit ganz zu Unrecht, so meinen wir, ein Stiefkind der Kognitionspsychologie. Dieses zeitweise getrübte Verhältnis zur Kognitionspsychologie offenbart sich in den Forschungsschwerpunkten im Bereich der Psychologie der auditiven Wahrnehmung und der Musikpsychologie bis ca. 197o.

Im Bereich der auditiven Wahrnehmung gibt es bis 197o die Schwerpunkte "Sprachwahrnehmung" und "Psychoakustik", wobei sich die Psychoakustik überwiegend mit den präattentiven Verarbeitungsmechanismen isolierter akustischer Signale beschäftigt. Die Ansätze, psychoakustische Befunde auf die Wahrnehmung von Musik anzuwenden, mußten solange unbefriedigend ausfallen, wie versucht wurde, Musikhören aus der Wahrnehmung isolierter physischer Merkmale von Tönen zu erklären (z.B. Risset, 1978; Ward, 197o). Dieses Defizit hat Hirsh (1967) gesehen und programmatisch gefordert, daß es für die Theorie der auditiven Wahrnehmung notwendig sei, eine Konzeption für zeitliche Muster zu entwickeln, die die gleiche Rolle für die Theorie der audi-

tiven Wahrnehmung spielen könnte, wie die Konzepte "Gestalt" und "Form" für die visuelle Wahrnehmung.

Die Ursache für die weitgehende Vernachlässigung der Probleme der Wahrnehmung von Musik sehen wir historisch begründet in der Trennung von Musikpsychologie und Psychologie der auditiven Wahrnehmung, die zur Ausbildung unterschiedlicher Arbeitsschwerpunkte führte und damit integrative Bestrebungen bereits im Keim erstickte.

Bei Hermann v. Helmholtz (1863) stand zwar noch der einzelne Ton als Tonempfindung im Mittelpunkt, aber das Ziel, auf einer Psychophysik der Tonempfindungen aufbauend eine Theorie des komplexeren musikalischen Hörens zu entwickeln, wurde bereits von ihm thematisiert.

In der Folgezeit stellte sich die Erkenntnis ein, daß eine Suche nach Elementen des musikalischen Hörens nicht sinnvoll sei, weil in der Introspektion beim Hören von Musik keine derartigen Elementarempfindungen zu beobachten waren. Obwohl für Stumpf (1883/189o) noch der Begriff der Elementarempfindung zentral blieb und er infolgedessen auch noch von einer Tonpsychologie sprach, versuchte er nicht mehr, wie dies noch bei Helmholtz (1863) der Fall war, die Beziehung zwischen den Elementarempfindungen durch eine additive Verknüpfung zu lösen. Zwar sei das Wahrnehmungsresultat in Elemente analytisch zerlegbar, aber spontane Eindrücke, wie sie sich beim Hören von Musik einstellen, erscheinen in der Introspektion des Hörers als etwas Einheitliches, das der Wahrnehmung direkt und unmittelbar gegeben sein müsse, ohne vorausgehende Verknüpfung von Elementen.

In der nachfolgenden Zeit entstand auf diesem Hintergrund die Gestalt- und Ganzheitspsychologie, die den Begriff der Elementarempfindung gänzlich fallen ließ. Bei Kurth (1931) taucht zum erstenmal der Begriff "Musikpsychologie" auf. Die Grundlage seiner Musikpsychologie bestand in der Erkenntnis, daß zwischen dem physikalisch-physiologisch beschreibbaren Reiz und seiner psychischen Wirkung keine starre Entsprechung in Form einer Abbildung besteht, wie dies noch für Helmholtz und Stumpf galt. Vielmehr müsse es Gegenstand einer Musikpsychologie im Unterschied zu einer Tonpsychologie sein, die Prinzipien der Wahrnehmungsorganisation zu untersuchen.

Was bei Kurth noch einem Wechsel der theoretischen Perspektive entsprach, wenn er den Begriff "Musikpsychologie" statt den der "Tonpsychologie" verwendet, wird z.B. bei Wellek (1963) zu einer Unterscheidung zwischen verschiedenen Forschungsgebieten, wodurch eine Trennung und isolierte Weiterentwicklung der Psychologie der auditiven Wahrnehmung bzw. der Psychoakustik und der Musikpsychologie manifestiert wurde. Damit einher ging die beinahe vollständige Vernachlässigung kognitiver Prozesse im Bereich der Musikpsychologie, deren Untersuchung als Gegenstand einer Tonpsychologie betrachtet wurde, aber die Tonpsychologie beschränkte sich gänzlich auf die Phänomene und Prozesse der Verarbeitung isolierter akustischer Signale, woraus - hauptsächlich in den USA - die Disziplin der Psychoakustik entstand.

Gleichzeitig mit dieser Trennung gewann der Behaviorismus auch unter Musikpsychologen vorwiegend in den USA an Ansehen, was zu einer vollständigen Eliminierung kognitiver Prozesse in musik-

psychologischen Theorien führte, wie dies anhand der behandelten Themen in der hierfür typischen Monographie von Lundin (1967) nachgewiesen werden kann.

Seitdem hat sich die Musikpsychologie hauptsächlich mit emotionalen Reaktionen auf Musik (z.B. Hevner, 1934, 1935, 1936, 1937; Capurso et al., 1952; Pratt, 1952; Rigg, 1964), mit experimenteller Musikästhetik (z.B. Hevner, 193o; Vitz, 1974), mit der Sozialpsychologie des musikalischen Geschmacks (Farnsworth, 1958), mit der Entwicklung musikalischen Verhaltens und der Diagnostik der Musikalität beschäftigt (vgl. Kleinen, 1972; Shuter, 1968). Aber selbst in diesen Arbeitsgebieten ist m.E. seit Ende des zweiten Weltkriegs eine Stagnation der Forschung zu beobachten. Dies liegt nicht etwa daran, daß alle musikpsychologischen Probleme hinreichend bearbeitet und alle Fragen zufriedenstellend beantwortet wären, sondern eher daran, daß die theoretische Reflektion und das methodische Repertoire der heutigen Musikpsychologie nicht mit den Entwicklungen in anderen Bereichen der Psychologie, insbesondere der Kognitionspsychologie, Schritt gehalten hat.

In der Psychologie der auditiven Wahrnehmung klafft ganz offensichtlich eine große Erkenntnislücke zwischen den peripheren Prozessen der Verarbeitung isolierter Töne (Domäne der Psychoakustik) einerseits und den emotionalen Prozessen im Erleben von Musik (Domäne der klassischen Musikpsychologie) andererseits.

Ein Forschungsprogramm zur Schließung dieser Lücke wurde von Deutsch (1969) initiiert, als sie ein Modell der kognitiven Ana-

lyseprozesse bei der Wahrnehmung von Musik entwarf. Der von Deutsch (1969, 1975a,b) und später zahlreichen anderen Autoren (z.B. Divenyi & Hirsh, 1973, 1974, 1975; Dowling, 1978; Handel, 1973, 1974; Hirsh, 1974; Massaro, 1975) verfolgte Ansatz versucht, ähnlich wie dies Neisser (1967) allgemein als Aufgabe der kognitiven Psychologie formulierte, das Schicksal wahrgenommener Töne und Tonfolgen auf ihrem Weg der Transformation, Speicherung und Abrufung aus verschiedenen Gedächtnissystemen zu verfolgen und Modelle zu entwickeln, die die Funktion und Struktur der beteiligten kognitiven Prozesse erklären können.

Jedoch ist es von der Psychologie der auditiven Informationsverarbeitung zu einer die kognitiven Prozesse des Musikhörens in den Mittelpunkt rückenden kognitiven Musikpsychologie noch ein weiter Weg. Während die traditionelle Musikpsychologie von der sich abzeichnenden Entwicklung unangefochten weiterlebt, hat sich die Kognitionspsychologie der Thematik des Musikhörens angenommen, ohne das Nebeneinander thematisch und historisch eng verknüpfter Forschungsfragen in der Musik- und Kognitionspsychologie damit aufheben zu können. Die kognitionspsychologische Forschung in diesem Bereich glaubt offenbar weiterhin, ohne Berücksichtigung der spezifisch musikalischen Strukturen eine Theorie der Musikwahrnehmung auf der Grundlage experimenteller Untersuchungen mit künstlich konstruierten, "unmusikalischen" Tonfolgen entwickeln zu können, und die musikpsychologische Forschung glaubt weiterhin, ohne Berücksichtigung kognitiver Prozesse auszukommen.

Pointiert formuliert können die bisher formulierten Modelle der auditiven Informationsverarbeitung lediglich Prozesse der Ent-

stehung kognitiver Repräsentationen von Musik erklären, wie sie bei einem Hörer ablaufen könnten, der noch nicht die geringsten Kenntnisse über musikalische Formen und Strukturen besitzt, denn ein eventuell beim Hörer implizit oder gar explizit vorhandenes Wissen (zum Begriff des "impliziten Wissens" vgl. Polanyi, 1966; zu seiner psychologischen Verwendung vgl. Turvey, 1974) über musikalische Strukturen wird nicht in seiner möglichen Funktion beim Hören von Musik thematisiert. Dies ist eine Konsequenz eines von der sensorischen Aufnahme bis hin zur langfristigen Speicherung linear verlaufenden Prozeßmodelles der auditiven Verarbeitung. Daß solche Modelle allgemein eine zu sehr vergröberte Darstellung der unter realen Wahrnehmungsbedingungen ablaufenden kognitiven Prozesse geben, hat Neisser (1976) überzeugend dargestellt.

Um das beim Hörer vorhandene Wissen in ein funktionales Modell des Musikhörens integrieren zu können, bedarf es einer Erweiterung der Forschungsstrategien im Rahmen einer kognitiven Musikpsychologie. Statt ausschließlich die Verarbeitung von der Reizseite aus zu betrachten und die Ausbildung unterschiedlicher Repräsentationsformen zu verfolgen, müßte zusätzlich die Funktion vorhandener Strukturrepräsentationen für die Prozesse der Reizmusteranalyse berücksichtigt werden. D.h. wir müssen die kognitive Verarbeitung musikalischer Strukturen als Interaktion zwischen aufsteigenden Prozessen der Reizanalyse und absteigenden Prozessen der Aktivierung und Nutzbarmachung von Repräsentationen über musikalische Strukturen auf unterschiedlichen Ebenen des Verarbeitungsprozesses begreifen. Damit wird das beim Hörer vorhandene Wissen über Musik zum zentralen Untersuchungsgegenstand.

Eine ähnliche Forschungsperspektive sieht auch Laske (1974, 1975, 1977b), wenn auch mit anderen Schwerpunkten:

"... the practitioners of traditional music theory have never raised the crucial question of what is the function of musical syntax within the framework of the total memory system involved in musical listening, composition, and reproduction." (Laske, 1977b, S. 8).

"A comprehensive theory of music is a science comprising two interrelated subdisciplines: first, a theory of musical structures, or musicology; and second, a theory of cognitive processes resulting in musical structures, or psychomusicology" (Laske, 1977b, S. 1o).

Aus letzterem Zitat geht unter anderem hervor, daß es Laske's Anliegen ist, die kognitiven Prozesse der musikalischen Performanz zu erforschen. M.E. geht diese Perspektive zum gegenwärtigen Zeitpunkt noch mindestens einen Schritt zu weit über den augenblicklichen Forschungsstand in diesem Bereich hinaus, wie auch die noch sehr unspezifischen theoretischen Ausführungen Laske's zu den beteiligten kognitiven Prozessen deutlich machen. Die naheliegendere Perspektive wird auch von Laske gesehen, aber von ihm selbst nicht verfolgt:

"By their very nature, music-syntactic hypotheses are manifestations of the esthetic pole of musical experience, that is, systematizations of musical perceptions. They ought to be verified by studies in musical thinking, especially in music-analytic perception" (Laske, 1977b, S. 1o9). Und an anderer Stelle: "Although the analogy between music and language is often used and the term 'musical grammar' appears here and there in the literature, the need for an equivalent of psycholinguistics in musical studies has remained curiously unperceived" (Laske, 1977b, S. 23).

Aufgabe einer kognitiven Musikpsychologie sollte es demnach sein, im Verhältnis zur Musikwissenschaft die Rolle zu übernehmen, die die Psycholinguistik für die Linguistik schon seit ca. 3o Jahren

besitzt, nämlich sich als diejenige Disziplin zu verstehen, die 1. die kognitive Adäquatheit musikalisch-syntaktischer Beschreibungen empirisch-psychologisch überprüft und 2. die beim Hören, Lernen und Behalten von Musik ablaufenden kognitiven Prozesse detailliert untersucht, um eventuell Anstöße zu kognitiv adäquateren musikalisch-syntaktischen Beschreibungen zu geben.

1.2 METHODEN DER FORMALEN STRUKTURBESCHREIBUNG VON MUSIK

Eine musikalische Strukturbeschreibung muß in zweierlei Hinsicht eine adäquate Beschreibung liefern. Sie muß zuerst einmal d e s k r i p t i v a d ä q u a t sein, d.h. einer musikalischen Struktur eindeutig eine bestimmte geordnete Menge von Regeln zuordnen. Doch mit deskriptiver Adäquatheit alleine kann sich allenfalls die Musikwissenschaft begnügen. Vielmehr ist es ja hier das Ziel, auf der Folie einer deskriptiv adäquaten Strukturbeschreibung eines musikalischen Formtyps Hypothesen über die Struktur der Repräsentation von Musik beim Hören abzuleiten und diese einer experimentellen Überprüfung zu unterziehen.

Diese Forschungsstrategie wurde bereits zur Zeit der Entstehung einer Musikpsychologie (damals noch "Tonpsychologie") von Riemann (1889) antizipiert. Riemann erhoffte sich von den Arbeiten seines Zeitgenossen Karl Stumpf eine psychologische Bestätigung seiner Theorie der Funktionalharmonik (de la Motte-Haber, 1976). Diese Hoffnung aber hat sich bis heute nicht erfüllt, nicht weil sich Riemann's Theorie als psychologisch wenig plausibel darstellen würde, sondern weil offenbar seit hundert Jahren kaum eine Kooperation oder gegenseitige Befruchtung zwischen Musikwissenschaft und Musikpsychologie stattgefunden hat, die zu einer experimentalpsychologischen Überprüfung musikwissenschaftlicher Theorien, soweit sie Aussagen über das menschliche Hören enthalten,

geführt haben könnte. Dieser Tatbestand muß aber um so mehr überraschen, als zumindest in Deutschland Musikpsychologie überwiegend von Musikwissenschaftlern und Musikpädagogen betrieben wird.

Watt (1973) sagt, für ihn gäbe es nur eine linguistische Realität und das sei die mentale Realität. Hiermit wendet sich Watt (1973) gegen die linguistischen Adäquatheitskriterien, die lediglich Prinzipien wissenschaftlicher Objektivität und Ökonomie der strukturellen Beschreibung sind. Deskriptive Adäquatheit ist gewiß eine notwendige Voraussetzung; aber in unserem Zusammenhang kann eine Strukturbeschreibung erst dann als adäquat bezeichnet werden, wenn sie über die deskriptive Adäquatheit hinaus in Einklang gebracht werden kann mit psychologischer Evidenz über die in der Strukturbeschreibung implizierten Hypothesen über die Repräsentation der beschriebenen Strukturen im Fall ihrer akustischen Realisierung und Wahrnehmung durch einen Hörer. Die von Watt (1973) geforderte "mentale Grammatik" der Sprache muß dann nicht nur im linguistischen Sinne deskriptiv adäquat sein, sondern auch k o g n i t i v a d ä q u a t . Kognitive Adäquatheit ist ebenso für eine musikalische Strukturbeschreibung zu fordern, wenn sie wie hier Grundlage für die Ableitung von Hypothesen über die Struktur der kognitiven Repräsentation sein soll.

Kognitive Adäquatheit läßt sich nur in einem iterativen Verfahren erreichen, indem bereits bei der ersten Formulierung einer Syntax das vorhandene Wissen über kognitive Prozesse so berücksichtigt wird, daß die aus der Syntax ableitbaren Hypothesen nicht im Widerspruch zum als "gesichert" geltenden Repertoire kognitionspsychologischer Befunde stehen. Daran anschließen muß sich ein Prozeß der Überprüfung neuer Hypothesen, dessen Ergebnisse eventuell

eine Neuformulierung der Syntax und eine erneute Phase der Überprüfung notwendig machen.

Wir haben unsere Syntax deutscher Volks- und Kinderlieder (s.Anh.1) bereits so formuliert, daß sie auf dem Hintergrund des gegenwärtigen Forschungsstandes als kognitiv adäquat zu bezeichnen ist (wir werden auf diesen Punkt gleich noch im Detail zurückkommen). Die noch zu berichtenden Experimente 3 bis 5 dienen einer Überprüfung ihrer kognitiven Adäquatheit bezüglich des Aspekts der Repräsentationsstruktur beim Hören von Melodien, die aufgrund dieser Syntax generiert wurden.

Die Versuche, Musik objektiv und formal strukturell zu beschreiben, sind bereits zahlreich, zeichnen sich aber alle - bis auf einige wenige linguistische Ansätze - dadurch aus, daß sie entweder schon nicht deskriptiv adäquat sind (Regelsysteme für serielle Muster, z.B. Deutsch (1980b), Restle (1970)), oder weder deskriptiv noch kognitiv adäquat sind (Informationstheorie).

1.2.1 INFORMATIONSTHEORIE DER MUSIK

Einer der ersten Versuche, musikalische Strukturen durch quantitative Methoden zu beschreiben, war die Anwendung der Informationstheorie Shannon's (Shannon & Weaver, 1949) durch Moles (1956). Mit dem mathematisch-statistischen Instrumentarium der Informationstheorie hat man in der Folgezeit stilistische Unterschiede zwischen Komponisten und musikhistorischen Epochen nachzuweisen

versucht (Bean, 1966; Brincker, 197o; Coons & Kraehenbuehl, 1957; Fucks, 1962, 1965, 1968; Fucks & Lauter, 1965; Hiller, 1964; Hiller & Fuller, 1967; Lewin, 1968; Stahl, 1964; Winckel, 1964; Youngblood, 1958), Methoden zur künstlichen Musikerzeugung erprobt (Pinkerton, 1956; vgl. auch Cohen, 1962), musikästhetisch-psychologische Untersuchungen über affektive Reaktionen auf Musik durchgeführt (vgl. Vitz, 1974; Werbig, 1971) und sie informationstheoretisch zu begründen versucht (Meyer, 1957; Moles, 1971).

Die informationstheoretische Behandlung musikalischer Ereignisfolgen beabsichtigt, ihre Strukturierungen quantitativ und objektiv erfaßbar zu machen. Die "exakt-wissenschaftliche Musikanalyse" (Fucks & Lauter, 1965) "kehrt die Spärlichkeit der Resultate eher hervor, als daß (sie) sie bemänteln könnte" (de la Motte-Haber, 1972, S. 92). Die Anwendungen der Informationstheorie auf musikalische Strukturen beschränkten sich überwiegend auf eine einzige musikalische Dimension, meist auf die Tonhöhen- und Intervallverläufe einer Stimme.

Die Diskrepanz zwischen dem aufgrund des Notentextes berechneten Redundanzmaß und der subjektiven Information des Hörers zeigte Francès (1958, S. 12o) am Beispiel eines Walzers von Chopin, für dessen harmonische Struktur sich ein äußerst hoher Redundanzwert ergab. Dieser Wert kam dadurch zustande, daß in dem ganzen Walzer die Subdominante fehlte. Für den Hörer, der in bestimmten musikalischen Kontexten die Subdominante erwartet, stellt ihr Fehlen einen großen subjektiven Neuigkeitswert dar, mit dem eigentlich eine niedrige Redundanz korrespondieren sollte.

Fucks' (1968, Fucks & Lauter, 1965) stilistische Analysen entlarven ihre Inadäquatheit besonders deutlich am Beispiel der 6. Sinfonie von Tschaikowsky, die aufgrund der berechneten Redundanz in dieselbe stilistische Gruppe fällt, wie Bergs Streichquartett Op. 3, Weberns Streichquartett Op. 28 und Bartoks 2. Sinfonie. Dieses Ergebnis fehlgeschlagener informationstheoretischer Quantifizierungsbemühungen ist darauf zurückzuführen, daß die schematische Auszählung der Tonhöhen der 1. Violinenstimme ähnliche Resultate erbringen mußte, wie sie bei Gleichverteilung aller zwölf Töne einer Tonleiter zu erwarten sind, weil in dieser Sinfonie die 1. Violine ein Motiv chromatisch sequenziert (de la Motte-Haber, 1972, S. 93-94); aber hierdurch ist noch lange nicht eine stilistische Nähe zur seriellen Musik gegeben.

Es gibt eine ganze Reihe von Gründen, den informationstheoretischen Ansatz einer Strukturbeschreibung von Musik als deskriptiv inadäquat zu qualifizieren. 1. Cohen (1962) hat in einer leider auf wenig Resonanz gestoßenen Arbeit nachgewiesen, daß die mathematischen Voraussetzungen zur Übertragung der Informationstheorie auf die Analyse musikalischer Strukturen nicht erfüllt sind. Z.B. muß das Kriterium der Ergodizität erfüllt sein, das besagt, daß die Wahrscheinlichkeit für das Auftreten eines Elements des Zeichenrepertoires zu jedem Zeitpunkt der Informationsübermittlung durch den Sender gleich sein muß. Ergodisch aber ist Musik nicht - außer in einigen Kompositionen von John Cage vielleicht -; so ist z.B. die Wahrscheinlichkeit für das Auftreten der Tonika auf der ersten Zählzeit eines Vordersatzes in einem deutschen Volkslied nahe an 1, während sie in der Mitte des Liedes deutlich geringer ist und am Ende des Nachsatzes sogar exakt den Wert 1

erreicht (vgl. die Syntax, Anh. 1).

Die Nichtergodizität von Musik ist eine Folge ihrer Gebundenheit an Formtypen. Allgemein formuliert heißt das, daß die informationstheoretischen Strukturbeschreibungen unsensibel gegenüber einer positionalen Gebundenheit einzelner musikalischer Elemente sind. Diese Charakterisierung trifft auch auf die informationstheoretische Analyse serieller Musik zu.

2. Bei ihrer Anwendung auf Musik wurde nicht hinreichend begründet, was ein "Zeichen" oder "Element" ist. Dies kann ein Ton, ein Takt oder ein Motiv sein. Die Willkür, die bei einer solchen Festlegung walten muß, trübt die "Exakt-wissenschaftlichkeit" der Ergebnisse.

3. Chomsky's (1956) Argumente, die begründen, warum Markoff-Quellen keine "self-embeded structures" (hierarchische Gliederung von Substrukturen) erzeugen können, gelten nicht nur für die Sprache, sondern ebenso für musikalische Strukturen, die häufig derartige hierarchische Organisationen enthalten.

4. Die informationstheoretische Analyse vernachlässigt die Kontextabhängigkeit des Auftretens einzelner musikalischer Elemente und die daraus resultierenden unterschiedlichen Funktionen eines Elements, z.B. kann die 5. Tonstufe auf unbetonter Zählzeit nur als harmonisch zur Dominante gehörig interpretiert werden, wenn ihr auf betonter Zählzeit ein anderer Ton der Dominante (2. oder 7. Tonstufe) vorausgeht, aber als zur Tonika gehörig, wenn ihr ein anderer Ton der Tonika vorausgeht (1. oder 3. Tonstufe).

Wenn eine formale Strukturbeschreibung schon nicht deskriptiv adäquat ist, kann sie kaum kognitiv adäquat sein. So setzt die Anwendung der Informationstheorie einen idealen Empfänger voraus, der über eine optimale und effiziente Dekodierungsstrategie verfügt, Kenntnis über das vollständige Zeichenrepertoire des Senders besitzt, alle Übergangswahrscheinlichkeiten beliebig hoher Ordnung für alle architektonischen Niveaus der Musik (Melodik, Harmonik, Rhythmik usw.) gespeichert hat und hierüber zu jedem Zeitpunkt in konstant bleibender Weise verfügen kann (Cohen, 1962). Aufmerksamkeitsfluktuationen, begrenzte Kapazität des kurzfristigen Behaltens und andere kognitive Funktionscharakteristika stehen der Annahme eines solchen Idealhörers entgegen.

Aus der Informationstheorie à la Moles (1971) und Meyer (1957) läßt sich die Vorhersage ableiten, daß jemand, der ein zuvor auswendig gelerntes Stück hört, dieses als langweilig empfinden müßte, weil der Hörer ja jeden einzelnen Ton richtig antizipieren kann, so daß die Musik für ihn keine Information mehr überträgt. Dies kann wohl im Ernst niemand so pauschal akzeptieren; obwohl die Literatur zur Veränderung ästhetischer Urteile bei mehrfacher Wiederholung von Musik diese Vorhersage teilweise bestätigt (Edmonston, 1969; Mull, 194o, 1957; Verveer, Barry & Bonsfield, 1933), zeigt doch die alltägliche Hörerfahrung, daß es auch das umgekehrte Phänomen gibt: Ein Musikstück, das man bereits häufig gehört hat, wird eigentlich erst hierdurch zum persönlichen Favoriten.

1.2.2 KODIERSYSTEME FÜR SERIELLE MUSTER

Die Anwendung der Informationstheorie in anderen Bereichen der

Psychologie (vgl. Attneave, 1965) stieß ebenso wie in der Musikpsychologie auf fundierte Kritik. Aus dieser Ablehnung der Informationstheorie entwickelten sich alternative Methoden zur Strukturbeschreibung, die zumindest implizit von einer psychologischen Äquivalenz der Strukturbeschreibungsmethoden und der kognitiven Prozesse des Strukturerkennens ausgingen.

Aus der Tradition der "artificial intelligence" - Forschung gingen die musikalischen Strukturerkennungsprogramme von Simon & Sumner (1968) und Smoliar (1971) hervor sowie die Theorie kompositorischer Problemlösungsprozesse von Laske (1977b), auf deren Grundlage Programme zur Darstellung einzelner Problemlösungsschritte und der dabei produzierten musikalischen Teilstrukturen entwickelt worden sind.

Diesen drei Ansätzen ist gemeinsam, daß sie musikalische Strukturen in einer computergerechten Form kodieren, die es gestattet, Computer zur musikalischen Strukturanalyse einzusetzen. Da die Kodierung von Element zu Element vorgenommen wird, ist ein Erkennen übergreifender und formgebundener Abhängigkeiten einzelner Substrukturen bei diesen Ansätzen nicht möglich. Zudem können nur individuelle musikalische Strukturen kodiert werden, so daß die einer Menge von Strukturen gemeinsamen Merkmale, die Inhalt eines musikalischen Wissens sein könnten, durch diese Methode nicht darstellbar sind. Damit verbunden ist die Unfähigkeit dieser Form der Strukturbeschreibung, ein Kriterium dafür zu liefern, ob eine vorliegende Tonstruktur innerhalb eines bestimmten musikalischen Formtyps und eines bestimmten Kompositionsstils eine Tonfolge ist, die noch zu diesem Formtypus und Kompositionsstil zu rechnen ist, oder bereits seine Grenzen sprengt.

Aus der Kritik an assoziationstheoretischen Analysen des seriellen Lernens entwickelten sich seit Mitte der 6o-iger Jahre flexiblere Theorien der kognitiven Kodierung serieller Reizfolgen. Starkes Interesse fanden nicht-lineare Strukturen, an denen sich die Unzulänglichkeiten assoziationstheoretischer Modelle besonders deutlich demonstrieren ließen (Deutsch, 1980b; Jones, 1978; Jones et al., 1978; Martin, 1972; Restle, 197o).

Deutsch (1980b), Restle (197o; Restle & Brown, 197o) und Jones (1974, 1978) verwandten dabei Regelsysteme, die hierarchische Strukturen gemeinsam mit einer Strukturbeschreibung generieren können, um diese zur Erzeugung von systematisch variierten Reizstrukturen zu verwenden, die in Experimenten zum seriellen Lernen eingesetzt wurden.

Sowohl Restle (197o) als auch Jones (1974, 1978) behaupten, daß ihre Regelsysteme zur Beschreibung musikalischer Strukturen geeignet seien und demonstrieren dies an einigen wenigen Beispielen. Es lassen sich aber ohne große Mühe zahllose Beispiele anführen, die mit keinem dieser Regelsysteme beschreibbar sind.

Bei den Regeln handelt es sich um sogenannte ordinale Transformationen eines Anfangselements durch einen Operator. Das Anfangselement ist Element einer geordneten Menge (z.B. der acht Töne einer diatonischen Tonleiter). So besagt z.B. die Regel: (R(1)) → (1,1), daß als Anfangselement das erste Element der Menge verwendet wird und auf dieses die Operation R (Wiederholung) angewendet werden soll, wodurch sich das rechts vom Pfeil stehende geordnete Paar ergibt. Mehrere Operatoren, wie z.B. Spiegelbildtransformation (M) und Transposition (T), lassen sich zu komplexen Regel-

strukturen kombinieren: z.B. $(M(T_{+2}(R(1))))$ → $(M(T_{+2}(1,1)))$ → $(M(1,1,3,3))$ → $(1,1,3,3,6,6,4,4)$ (die M-Operation besagt, daß das letzte Element, also das achte, zum Anfangselement der Entwicklung wird und die Transpositionsrichtung von +2 nach -2 umgekehrt wird).

Da dieses Regelsystem ausschließlich symmetrische Strukturen erzeugen und beschreiben kann, Musik aber durchaus nicht nur derartig symmetrische Strukturen enthält, kann dieser Ansatz nicht als deskriptiv adäquat zur Beschreibung musikalischer Strukturen bezeichnet werden. Obwohl die oben zitierten experimentellen Arbeiten von Restle und Jones nachgewiesen haben, daß das Verhalten der Vpn (Antizipationsfehler beim Lernen der seriellen Muster) von der Regelstruktur in starkem Maße abhängig ist, kann das Gesamtsystem auch nicht als für die Wahrnehmung musikalischer Strukturen kognitiv adäquat bezeichnet werden, weil es ihm schon an deskriptiver Adäquatheit mangelt. Der Ertrag dieser Forschungen ist aber dennoch für eine Theorie der Musikwahrnehmung von Bedeutung, weil nachgewiesen wurde, daß auch nicht-sprachliche Muster kognitiv als hierarchische Struktur repräsentiert werden können.

1.2.3 ÜBERTRAGUNG LINGUISTISCHER METHODEN AUF DIE ANALYSE UND BESCHREIBUNG MUSIKALISCHER STRUKTUREN

In den letzten zwanzig Jahren sind vereinzelt Stimmen laut geworden, die eine Einbeziehung linguistischer Methoden in das Repertoire musikwissenschaftlicher Analysetechniken propagieren

(Nattiez, 1972, 1973; Nettl, 1958; Ruwet, 1972; Springer, 1956). Wenn man zurückverfolgt, welche Impulse die Psycholinguistik von der Theorie der Transformationsgrammatik Chomsky's (1957, 1965) empfangen hat, dann kann man hoffen, daß die Musikpsychologie eine ebenso geartete Stimulierung erfährt, wenn es gelingt, musikalische Strukturen zunächst erst einmal deskriptiv adäquat und formal darzustellen. Mit der Realisierung eines dem wissenschaftlichen Vorgehen der modernen Psycholinguistik analogen Forschungsprogramms im Rahmen einer kognitiven Musikpsychologie verbinden wir die Hoffnung, allmählich auch zu einer kognitiv adäquaten Beschreibung musikalischer Strukturen zu gelangen, die als Hypothese der Repräsentation musikalischer Strukturen betrachtet werden kann. Diese Forschungsperspektive kann sich umso enger an die der Psycholinguistik anlehnen, je enger und plausibler die strukturellen Analogien zwischen Sprache und Musik sind.

Analogien zwischen Sprache und Musik

"Die Sprache der Musik verfügt über alle Elemente, die es nicht nur erlauben, eine Analogie zu den natürlichen Sprachen herzustellen, sondern sie in ähnlicher Weise zu analysieren" meint Schaff (1973, S. 281). Dem stimmen die meisten Autoren, die sich mit der Semiotik der Musik beschäftigt haben, nicht uneingeschränkt zu.

Bezüglich des syntaktischen Aspekts der musikalischen Strukturen sind sich alle einig: Musik kann als System regelhaft aufeinander bezogener Einheiten aufgefaßt werden (Adorno, 1963; Bright, 1963; Faltin, 1973; Harweg, 1967, 1968, 1972; Koch, 1971; Nattiez, 1972,

1973; Nettl, 1958; Osmond-Smith, 1974; Ruwet, 1972; Schaff, 1973; Springer, 1956). Kontroverse Diskussionen über die Tragfähigkeit von Analogien zwischen Sprache und Musik entzündeten sich aber immer wieder an der Frage, ob Musik, wie die natürliche Sprache, ein System von Zeichen ist, das auch einen semantischen Aspekt besitzt.

Harweg (1967, 1968, 1972) sieht Bedeutung (im Sinne von Referenz) als konstitutiv für die Existenz eines Zeichens an, andere wie Bright (1963), Springer (1956) und Meyer (1956) sehen es als hinreichend an, wenn ein Element auf seine Stellung innerhalb eines Systems verweist, d.h. regelhafte Beziehungen zu anderen Elementen eingeht (vgl. Osmond-Smith, 1974).

Im streng semiotischen Sinne stellen Folgen musikalischer Elemente keine Zeichenfolgen dar und insofern ist Musik auch kein "Verstehensobjekt" (Harweg, 1972), obwohl es so etwas wie ein Wiedererkennen der Außenwelt in Musik gibt (z.B. das Blöken der Schafe in Richard Strauß'ens "Don Quichotte") und musikalische Zitate (z.B. das Muttersehnsuchtsmotiv der Chrysothemis aus der Oper "Elektra" in thematischer Rekurrenz in der Liebesmelodie des Kaisers in der Oper "Die Frau ohne Schatten" von Richard Strauß, vgl. Harweg et al., 1967). Bei diesen Fällen handelt es sich aber um Beispiele für die relativ eingeschränkten Möglichkeiten, die Außenwelt nicht-arbiträr zu wiederholen (Koch, 1971).

Dieser Sachverhalt wirft einige Probleme für die syntaktische Beschreibung von Musik auf, wenn man sich an den semantisch orientierten Grammatiken für natürliche Sprachen anzulehnen versucht,

da bei Fehlen einer semantischen Komponente ein wichtiges Entscheidungskriterium für die Beurteilung der grammatikalischen Akzeptabilität eines (musikalischen) Satzes (Phrase oder Periode) fortfällt. Eine Transformationssyntax der Musik im Sinne der Grammatikmodelle von Chomsky (1957, 1965) kann es aus diesem Grunde nicht geben. Musik ist deshalb nur als formale Sprache zu betrachten.

Die syntaktischen Analogien zwischen Sprache und Musik dagegen sind sehr vielfältig, wenn man einmal von der Musik der Gegenwart absieht, deren syntaktische Organisation im Vergleich zur Musik zwischen Barock und Impressionismus relativ schwach ausgeprägt ist.

Der einzige deutlich durch ein syntaktisches System dominierte musikalische Parameter ist die Tonhöhe (Osmond-Smith, 1974). Musik ist aus einer endlichen Menge von Tonhöhen gebildet, die die minimalen und zugleich terminalen Elemente einer formalen Syntax darstellen können. Dabei ist es irrelevant zu fragen, wie dies Nettl (1958) tut, welche musikalischen Tongebilde welchen sprachlichen Einheiten entsprechen könnten; dies wäre nur dann eine sinnvolle Fragestellung, wenn man Musik wie eine natürliche Sprache behandeln wollte. Vielmehr ist es sinnvoll zu untersuchen, wie sich einzelne minimale Elemente zu Tongruppen höherer Ordnung zusammenschließen und welchen Regeln dieser Zusammenschluß unterliegt, d.h. es ist zu untersuchen, ob es eine der sprachlichen Struktur vergleichbare Konstituentenstruktur innerhalb bestimmter musikalischer Formen gibt und wie diese Konstituentenstruktur als durch Regeln ableitbar gedacht werden kann.

Außerdem ist zu untersuchen, und dies ist für die Musik einzigartig, ob es bestimmte Formen ordinaler Transformationen gibt und wie diese geregelt sind, z.B. musikalische Sequenzen und Wiederholungen (vgl. Ruwet, 1972). Diese Transformationen stellen keine durch ein linguistisches Ökonomieprinzip diktierten Rekursionen unterschiedlicher terminaler Strukturen auf eine ihnen gemeinsame Basis dar, wie dies bei der generativen Grammatik der Fall ist, sondern sie sind geregelte Veränderungen einer Tonfolge im Sinne mathematischer Funktionen. So ließe sich die musikalische Sequenzbildung als eine Art Parallelverschiebung um einen bestimmten Intervallbetrag darstellen. Diese Form der Transformationen ist für die Musik spezifisch, weil das Elementenrepertoire der Tonhöhen (z.B. die diatonische Tonleiter) zumindest eine ordinal geordnete Menge darstellt. Musik, als formale Sprache betrachtet, ist offenbar einer Analyse mit Mitteln der Syntaxtheorie im Prinzip zugänglich.

Musiktheoretische Anknüpfungspunkte

Das Verdienst, auf die Bedeutung der "musikalischen Logik" Riemann's (1889) für die Musikpsychologie hingewiesen zu haben, kommt de la Motte-Haber (1976) zu. Riemann's Beschreibungskategorien für Musik sind Kategorien der Wahrnehmung, für die Beziehungen zwischen Tönen und Tongruppen und die sie stiftenden Regeln von fundamentaler Bedeutung sind. Er beschreibt Musik mittels Relationsbegriffen, wie sie ebenfalls für die Beschreibung sprachlicher Strukturen adäquat sind.

Zwei Aspekte der Theorie Riemann's sind für uns hier von zentraler Bedeutung: 1. seine Funktionalharmonik und 2. sein Modell der musikalischen Periode. Beide Aspekte zusammen gestatten eine strukturelle Beschreibung von größeren wie von kleineren musikalischen Einheiten, die aufeinander regelhaft bezogen sind.

Nach Riemann ist das Hören von Akkorden oder auch das einzelner Töne einer unbegleiteten Melodie ein Hören von Beziehungen zu einem Hauptklang, nämlich der Tonika, die damit zum Bezugssystem für alle Töne wird. Neben der Tonika gibt es nur noch zwei weitere Bezugsfunktionen, nämlich die Dominante und die Subdominante. Durch die Funktion eines Tones als zur Tonika, Dominante, Subdominante oder zu der Klasse der harmoniefremden Töne gehörig werden Beziehungen und auch Abgenzungen zwischen benachbarten Tönen hergestellt. Damit kann die Vielfalt der Töne auf ganz wenige Kategorien reduziert werden, was für eine syntaktische Analyse, die im Prinzip eine Klassifikation auf verschiedenen Ebenen darstellt, unbedingt notwendig ist.

Eine vergleichbare Reduktion der musikalischen Vielfalt auf eine prototypische Kategorie stellt Riemann's Modell der musikalischen Periode dar. Als Norm einer Taktgruppierung werden danach acht Takte betrachtet und Abweichungen von diesem Schema als Auslassungen bzw. Hinzufügungen verstanden. Mehr noch als durch die Zahl der Takte wird eine Periode durch den akkordischen Verlauf gegliedert und von anderen Perioden abgrenzbar: Typisch ist der Halbschluß im vierten Takt auf der Quinte der diatonischen Tonleiter und der Ganzschluß auf der Tonika im achten Takt. Diese Untergliederungen der Periode werden "Vordersatz" (Takte 1 bis 4) und

"Nachsatz" (Takte 5 bis 8) genannt, wobei noch ein "Mittelsatz" hinzutreten kann, der die achttaktige Periode auf zwölf Takte erweitert. Wiederholungen und Sequenzbildungen von einzelnen Takten oder Taktgruppen können darüber hinaus die Periode noch erweitern.

Abraham & Dahlhaus (1972) bezeichnen die Prinzipien der Periodenunterteilung in eine Hierarchie immer kleiner werdender Einheiten als "rhythmisch-syntaktische Quadratur". Dabei schließen sich nach dem Prinzip der korrespondierenden Variante bzw. des ergänzenden Kontrasts regelmäßig zwei Takte zu einem Motiv bzw. zu einem dem Motiv folgenden Anschlußglied zusammen, Motiv und Anschlußglied bilden dann einen Satz, den Vorder-, Mittel- oder Nachsatz, und die Sätze schließlich bilden die Periode. Diese Periodenunterteilung ist zumindest für das Volks- und Kinderlied deutlich ausgeprägt und leicht nachweisbar (Ziegenrücker, 1979). Riemann's auf Harmonik und Metrik fußendes Modell der Taktgruppierungen läßt sich als eine Reihe von Ersetzungsregeln zur Ableitung musikalischer Perioden darstellen (vgl. Anh. 1).

Ein zweiter Anknüpfungspunkt an musiktheoretische Konzeptionen besteht in dem Konzept des "Ursatzes" von Schenker (1954). Ein Ursatz ist ein melodisch-harmonischer Kern, der durch Fortlassen von Durchgangsnoten abstrahiert wird und dessen "Verwandlungen" innerhalb eines Musikstücks in der Analyse verfolgt werden. Dieser durch Analyse des "Vordergrundes" freigelegte "Hintergrund" und seine "Verwandlungen" wurden von Anhängern Schenker's mit der Grammatiktheorie Chomsky's in Beziehung gebracht, wobei der "Vordergrund" als "Oberflächenstruktur", der "Hintergrund" als "Tiefenstruktur" und die "Verwandlungen" als "Transformationen" interpretiert wurden

(vgl. de la Motte-Haber, 1976).

Die oberflächliche Analogie des Begriffs "Verwandlung" zum Transformationsbegriff bei Chomsky beruht allerdings auf einem Mißverständnis. Transformationen beziehen unterschiedliche Oberflächenstrukturen auf eine ihnen gemeinsame Tiefenstruktur, wobei eine Bedeutungsgleichheit impliziert wird. Diese aber läßt sich bei Musik nicht festmachen, da Musik nicht ein System von Zeichen mit denotativer Bedeutung darstellt.

Trotz der Oberflächlichkeiten dieses Übertragungsversuchs scheint uns der Ansatz Schenker's einen durchaus richtigen Gedanken zu enthalten, nämlich den, daß es in Musik so etwas wie systematische Abwandlungen einzelner musikalischer Passagen gibt, die sich in Form von Regeln mit ordinalen Transformationen darstellen lassen (vgl. Anh. 1).

Erste Ansätze zur Formulierung musikalisch-syntaktischer Regeln

Eine Syntax ist ganz allgemein ein formales System zur Festlegung von Symbolfolgen über einem Vokabular. Dieses formale System besteht im wesentlichen aus Regeln, mit deren Hilfe Symbolfolgen durch eine endliche Anzahl von Regelanwendungen abgeleitet werden können (Kratzer et al., 1974). Sie stellt einen Algorithmus dar, der so explizit formuliert ist, daß er im Prinzip von einer Maschine ausgeführt werden könnte.

Wenn die Linguistik die ein Verhalten ermöglichende Kompetenz von

Sprachbenutzern beschreibt, geht sie rein deskriptiv vor. Insofern sie aber Regeln formuliert, mit denen sprachliche Texte erzeugt werden können, wählt sie eine generative Darstellungsweise. Die Schule der "generativen, transformationellen Grammatik", die auch kurz als "generative Grammatik" bezeichnet wird, stellt darüber hinaus ganz bestimmte Anforderungen an die zu verwendenden Regeltypen. Danach muß eine Grammatik aus Ersetzungs- und Transformationsregeln aufgebaut sein (Bünting, 1972).

Seit ca. zehn Jahren wird insbesondere auf dem Gebiet der Vergleichenden Musikwissenschaft in den USA, aber auch in Europa eine Darstellung musikalischer Strukturen mit den Mitteln der Linguistik versucht, was in einigen Arbeiten auch bis zur expliziten Regelformulierung durchgehalten werden konnte (Arom, 197o; Baroni & Jacoboni, 1978; Boiles, 1967; Chenoweth & Bee, 1971; Cooper, 1973, 1977; Lindblom & Sundberg, 197o; Levi-Strauss, 1971; Sapir, 1969; Sundberg & Lindblom, 1976; Winograd, 1968; weitere Arbeiten s. Stoffer & Boroda, 198o).

Einer dieser Ansätze, nämlich der von Sundberg & Lindblom (1976), erscheint für unsere Zwecke besonders interessant, weil dort schwedische Volks- und Kinderlieder durch eine generative Syntax beschrieben werden und wir für unsere Untersuchung denselben Formtypus heranziehen wollen. Deshalb wollen wir die in dieser Arbeit verwendeten Prinzipien der syntaktischen Beschreibung daraufhin kritisch untersuchen, ob sie für deutsche Volks- und Kinderlieder eine deskriptiv adäquate strukturelle Beschreibung ermöglichen.

Den Ausgangspunkt bildet eine Distributionsanalyse von harmonischen Funktionen und Typen harmoniefremder Töne sowie von Tondauern. Eine Distributionsanalyse sollte eigentlich zeigen, in welchen Kontexten eine bestimmte Kette syntaktischer Elemente vorkommt und welche anderen Ketten in demselben Kontext erlaubt sind, so daß die in denselben Kontexten vorkommenden Ketten zu einer Konstituentenklasse zusammengefaßt werden können. Bei Sundberg & Lindblom besteht der berücksichtigte Kontext ausschließlich in der Position eines Tones innerhalb der achttaktigen Periode, d.h. der tonale Kontext bleibt unberücksichtigt. Dieses Vorgehen hat zwar den "Vorteil", daß die untersuchten Melodiestrukturen als kontextfreie Syntax darstellbar werden, handelt sich aber dadurch auch eine geringe Sensibilität für weitgespannte Abhängigkeiten von Melodieparametern ein.

Unberücksichtigt bleiben auch einzelne Tonhöhen (Stufen innerhalb der diatonischen Tonleiter), da lediglich harmonische Funktionen berücksichtigt werden, so daß es fast an Zauberei grenzt, wenn bei der generativen Anwendung der Syntax tatsächlich alle generierten Melodien auf dem Grundton enden, obwohl die Syntax dies an keiner Stelle explizit vorschreibt.

Der nächste Schritt besteht in der Ableitung von Betonungsrängen aus der Klammerdarstellung der Konstituentenhierarchie nach Chomsky & Halle (1968). Dieses Verfahren setzt die richtige Segmentierung eines Satzes bzw. hier einer musikalischen Periode voraus, was aber nur durch detaillierte Distributionsanalysen möglich wird, die distributive Kriterien zur Abgrenzung von Segmenten liefern. Daß Sundberg & Lindblom (1976) keine zumindest

eindeutigen Kriterien an der Hand hatten, zeigt die falsche Konstituentenhierarchie (S. 117) der untersuchten Melodien (S. 114). Aufgrund des dargestellten Baumgraphen ist anzunehmen, daß der Nachsatz (Takte 5 bis 1o) aus drei Subphrasen besteht, die von einem ihnen gemeinsamen Knoten (Nachsatz) dominiert werden. Betrachtet man aber die ersten drei Melodien (S. 114), dann fällt auf, daß die Takte 7 und 8 eine Wiederholung der Takte 5 und 6 darstellen. Die diesem Sachverhalt gerecht werdende Struktur erfordert einen Baumgraphen, dessen Knoten "Nachsatz" sich nicht in drei auf derselben Hierarchiestufe stehende Subphrasen verzweigt, sondern in nur zwei Verzweigungen, nämlich Subphrase 1 und Subphrase 2, wobei der Knoten "Subphrase 1" sich wiederum in zwei Subphrasen aufspalten muß, nämlich in die gleichgebildeten Takte 5 und 6 sowie die Takte 7 und 8.

Die Ableitung der Betonungsränge geht von der Klammerdarstellung der Konstituentenhierarchie aus, wobei jede Zählzeit zunächst den gleichen Betonungsrang, nämlich Rang 1, erhält. Bei Auflösung der innersten Klammern wird das "left priority rule", bei allen weiteren das "right priority rule" angewendet. Beim "left priority rule" wird der links in der verbleibenden Klammer stehende Rang beibehalten, während alle rechts innerhalb derselben Klammer stehenden Ränge um 1 erhöht werden.

Dies führt dazu, daß der niedrigste Betonungsrang, der die stärkste Betonung symbolisiert, im 4/4 - Takt auf die dritte Zählzeit des letzten Taktes fällt. Diese Endbetonung mag zwar für die untersuchten schwedischen Lieder zutreffend sein, sie ist es aber keineswegs für alle deutschen Volks- und Kinderlieder. In Abhängig-

keit von den Betonungsrängen werden dann die Harmonien ausgewählt, wobei der Betonungsrang 1 immer mit der Tonika besetzt wird. In der Mehrzahl der deutschen Volks- und Kinderlieder erscheint die Tonika aber bereits auf der ersten Zählzeit des letzten Taktes der Periode. Würde man aber statt des "left priority rule" das "right priority rule" anwenden, dann wäre zwar der niedrigste Betonungsrang der auf der ersten Zählzeit des letzten Taktes, aber gleichzeitig wäre der Betonungsrang 2 der auf der ersten Zählzeit des letzten Taktes des Vordersatzes, was nicht unbedingt mit den in deutschen Volks- und Kinderliedern anzutreffenden Endungen des Vordersatzes übereinzustimmen braucht.

Unabhängig von diesen Problemen ergibt sich ein schwerwiegender Einwand gegen diese Ableitungsmethode daraus, daß ja sehr viele Lieder mit Auftakt beginnen, ohne daß dadurch die Konstituentenhierarchie eine andere Klammerdarstellung erhalten müßte. Durch den Auftakt aber verschieben sich die tatsächlichen Betonungen innerhalb einer Konstituentenklasse um eine Zählzeit, was bei Ableitung der Betonungsränge aus der Klammerdarstellung der Konstituentenstruktur aber nicht zu erzeugen ist.

In der Arbeit von Sundberg & Lindblom (1976) zeigt sich, daß eine Übernahme von Darstellungs- und Ableitungsmethoden, wie sie in der generativen Linguistik verwendet werden, nicht immer auf musikalische Strukturen in gleicher Weise anwendbar sind. Dies liegt zum einen darin begründet, daß Musik bedeutend variantenreicher als eine natürliche Sprache ist, zum anderen darin, daß die Analogie zwischen sprachlichen Strukturen und musikalischen Strukturen zwar an zahlreichen Beispielen nachzuweisen ist, daß

sich allerdings auch beinahe ebenso viele Gegenbeispiele finden lassen, wenn man nur eine genügend große Zahl musikalischer Strukturen untersucht. Aus diesem Grunde wird sich wohl nie eine allgemeine Syntax für a l l e musikalischen Formen und Kompositionsstile aufstellen lassen, vielmehr wird man gezwungen sein, für jede einigermaßen objektiv abgrenzbare musikalische Form und für die verschiedenen Kompositionsstile eigene Syntaxen zu entwickeln, denen zwar bestimmte Darstellungstechniken, insbesondere die verwendeten Regeltypen, teilweise gemeinsam sein können, nicht aber die spezifischen Inhalte ihrer Regeln.

Die Ableitung musikalischer Strukturen aus geklammerten Betonungsröngen erweist sich hier als nicht verallgemeinerungsfähige Methode. Leider sind die Autoren auch nicht ihrem Vorsatz treu geblieben, die Regeln zur Ableitung in expliziter Weise darzustellen, obwohl dies, wie wir in Anh. 1 gezeigt haben, möglich ist.

1.3 DESKRIPTIVE UND KOGNITIVE ADÄQUATHEIT DER SYNTAX DEUTSCHER VOLKS- UND KINDERLIEDER

Im Kontext der experimentellen Untersuchung musikpsychologischer Fragestellungen besitzt eine musikalische Syntax drei Funktionen: Sie ist ein explizit formuliertes Beschreibungssystem für die Struktur der notierten Musik, das als Basis zur Ableitung von Hypothesen über den Zusammenhang von musikalischer Struktur und Verhaltensdaten der diese Struktur wahrnehmenden Hörer dient; sie ist außerdem ein methodisches Instrument zur Herstellung von Musik als Reizmaterial, das es in objektiver Weise gestattet, strukturell

gleichartiges oder systematisch variiertes Reizmaterial zu erzeugen, so daß die strukturelle Vergleichbarkeit des Reizmaterials zwischen Versuchsbedingungen auch bei Darbietung völlig verschiedener Musik gewährleistet ist; und sie ist schließlich wegen der Eindeutigkeit ihrer begrifflichen Festlegungen Kommunikationsmittel zwischen Experimentator und Vpn, wenn den Vpn die syntaktischen Kenntnisse vermittelt wurden, die sie zur Lösung ihrer experimentellen Aufgaben benötigen.

Im folgenden wollen wir die Syntax kurz beschreiben (der an Details interessierte Leser sei auf den Anh. 1 verwiesen [1]) und anschliessend ihre deskriptive, insbesondere aber ihre kognitive Adäquatheit diskutieren, soweit dies aufgrund des Hierarchiemodells und vorliegender empirischer Befunde bereits möglich ist.

Beschreibung der Syntax

Der Syntax liegt ein Corpus von 71 einsätzigen deutschen Volks- und Kinderliedern zugrunde, wie sie bei Klusen (1978), Lutz (1950) und Schüngeler (o.J.) notiert sind.

Die musikalische P e r i o d e (Umfang einer ganzen Liedstrophe) besteht entweder aus acht, zehn, zwölf oder sechzehn Takten und ist entweder zwei- oder dreiteilig. Zweiteilige Perioden bestehen aus

---

1) Die in Anh. 1 dargestellte Syntax weicht in einigen Punkten von der bei Stoffer (1979) dargestellten Version ab. Diese Abweichungen sind durch Fehlerkorrekturen in der formalen Darstellung einiger Regeln begründet, die uns bei der ersten Formulierung unterlaufen waren.

zwei P h r a s e n (in der Regel vier Takte, die durch Wiederholungen auf bis zu acht Takte erweitert werden können), dreiteilige Perioden aus drei Phrasen (Vorder-, Mittel- und Nachsatz, wobei der Vordersatz oder der Nachsatz wiederholt sein kann). Jede Phrase gliedert sich in zwei S u b p h r a s e n (Motiv und Anschlußglied) von jeweils zwei Takten, kann aber durch Motivsequenz (Wiederholung des Motivs auf einer anderen Tonstufe der diatonischen Tonleiter) oder Anschlußgliedwiederholung erweitert sein. Der Ebene der Subphrasen ist die der T a k t e untergeordnet, die jeweils aus ein bis zwei F Ü ß e n bestehen (im 3/4-Takt ein Fuß mit Betonung auf dem ersten Taktschlag, im 4/4-Takt zwei Füße, wobei der erste Taktschlag des ersten Fußes die Hauptbetonung und der erste Taktschlag des zweiten Fußes die Nebenbetonung trägt). Der Ebene der Füße ist die der B e t o n u n g s r ä n g e untergeordnet. Jeder Taktschlag erhält aufgrund des Gewichts seiner Betonung einen Betonungsrang, dessen Größe umgekehrt proportional der Betonungsstärke ist, z.B. im 4/4-Takt gibt es die Ränge 1-3-2-3 für die einzelnen Viertelschläge (Abb. 1 in Anh. 1 veranschaulicht diese Periodengliederung durch einen Ausschnitt aus einem Baumgraphen).

Diese Periodenstruktur wird beginnend mit dem höchsten Hierarchieniveau ebenenweise durch Anwendung von Regeln erzeugt, die eine Einheit einer Ebene durch zwei oder drei Einheiten der unmittelbar untergeordneten Ebene ersetzen und eventuell bestehende Transformationsbeziehungen zwischen Einheiten einer Ebene spezifizieren. In der vertikalen Richtung der strukturellen Hierarchie wird also die jeweilige Binnengliederung der horizontal benachbarten Einheiten definiert, in der horizontalen Richtung wird gleichzeitig die strukturelle Verwandtschaft benachbarter Einheiten definiert, die auf

der übergeordneten Ebene zur selben Einheit gehören. Im Unterschied zu Chomsky's (1965) Transformationsregeln, die erst nach Ableitung einer Basisstruktur auf diese angewandt werden, kann in unserer Syntax eine Transformation auf jeder Ebene vorgenommen werden.

Beispiele für solche "ordinalen Transformationen" in "transformationellen Ersetzungsregeln" sind die W i e d e r h o l u n g z.B. eines Taktes oder eines Taktpaares, die S e q u e n z (Wiederholung der Tonfolge auf einem höheren oder tieferen Ton der diatonischen Tonleiter) und die U m k e h r u n g der Intervallrichtungen innerhalb einer Einheit, die gleichzeitig mit der Sequenz kombiniert sein kann. Dabei wird der Struktur einer Einheit A die transformierte Struktur A' rechts hinzugefügt, so daß auf einer Ebene die horizontale Gliederung AA' entsteht, wobei sowohl A als auch A' gemeinsame Elemente einer übergeordneten Einheit sein müssen.

Fast alle syntaktischen Regeln enthalten Kontextangaben. Diese definieren innerhalb des durch die hierarchische Periodengliederung festgelegten Rahmens lineare Abhängigkeiten, die bei der Auswahl zwischen alternativen Ersetzungsregeln bei der Erzeugung eines bestimmten hierarchischen Niveaus berücksichtigt werden müssen. Eine häufige Form solcher linearen Abhängigkeiten ist die K a d e n z r e g e l , die eine bestimmte Folge von Harmonien an bestimmten Stellen einer Periode vorschreibt bzw. an anderen Stellen ausschließt, oder die Regel, daß die Periode mit einem Ton der Tonika beginnen und enden muß.

Die terminalen Elemente der Syntax sind einzelne Töne mit ihren harmonischen Funktionen, die innerhalb eines bestimmten Kontextes "erlaubt" sind. Durchläuft man die Regelhierarchie von oben nach unten, dann kann man formal richtige "neue Volkslieder" generieren, durchläuft man sie in der umgekehrten Richtung, kann man die Regelstruktur eines notierten Volksliedes analysieren.

Deskriptive Adäquatheit der Syntax

Eine musikalische Syntax ist analog zum linguistischen Adäquatheitskriterium dann für eine definierte Teilmenge musikalischer Strukturen als eindeutig deskriptiv adäquat zu bezeichnen, wenn sie einer musikalischen Struktur eine Menge von Regeln in eindeutiger Weise zuordnet (analytische Adäquatheit) oder wenn sie einer Menge von Regeln genau eine musikalische Struktur zuordnet (generative Adäquatheit). Sie ist dann eineindeutig deskriptiv adäquat, wenn sie sowohl analytisch als auch generativ adäquat ist.

Eine Syntax zur Erzeugung syntaktisch richtiger Melodien, wie wir sie für die Experimente 3 bis 5 benötigen, muß mindestens generativ adäquat sein. Die Beurteilung der generativen Adäquatheit stellt für musikalische Syntaxen ein besonders schwieriges Problem dar. Während es bei natürlichen Sprachen das Kriterium der Sinnhaftigkeit eines generierten Satzes gibt, existiert für musikalische Syntaxen kein vergleichbares Kriterium mangels semantischer Komponente. Die Verallgemeinerung einer musikalischen Syntax über den Bereich des analysierten Corpus hinaus

kann sich zunächst nur auf die musikalisch-syntaktische Intuition desjenigen stützen, der diese Syntax formuliert.

Die empirische Überprüfung an musikalischem Material kann nur so weit gehen, daß nachgewiesen wird, daß die Syntax alle zum selben musikalischen Formtyp gehörigen Strukturen erzeugen kann. Da es aber nicht nur das Ziel der syntaktischen Beschreibung sein kann, die bereits existierenden Strukturen zu beschreiben, sondern es vielmehr auch das Ziel sein muß, alle potentiell zum selben Formtyp gehörigen Strukturen zu beschreiben, auch wenn diese erst zu einem späteren Zeitpunkt komponiert werden sollten, stellt die Überprüfung der generativen Adäquatheit unserer Meinung nach ein psychologisches Problem dar. Die Frage, ob eine musikalische Struktur, die von einer über den zugrundeliegenden Corpus hinaus verallgemeinerten Syntax generiert wurde, zum selben musikalischen Formtyp gehört, wie die im Corpus enthaltenen Strukturen, kann nur durch Befragung kompetenter Hörer beantwortet werden, deren Entscheidung sich auf ihr implizites oder explizites musikalisches Wissen stützt.

Da die Syntax aber die Basis zur Ableitung von Hypothesen darstellen soll, die nicht nur Aussagen über die Struktur der Wahrnehmungsrepräsentation sondern auch über die Struktur der Repräsentation eines musikalisch-syntaktischen Wissens machen sollen, die Hypothesenüberprüfung sich aber wiederum auf dieses Wissen stützen muß, ist es notwendig, die diesem Wissen und seiner Anwendung zugrundeliegenden kognitiven Prozesse zu kennen, die wir aber nur mit Hilfe gerade dieser Syntax erforschen können.

Die hier aufgewiesene Zirkularität bei der Beurteilung der generativen Adäquatheit einer musikalischen Syntax ist nur in der Hoffnung auf eine allmähliche Konvergenz zwischen der Adäquatheit unserer Kenntnisse über die kognitiven Prozesse beim Hören von Musik und der Adäquatheit der syntaktischen Beschreibung zu bewältigen. D.h. zunächst muß es einmal genügen, eine adäquate Syntax zu besitzen, deren generative Adäquatheit sich ausschließlich auf den zugrundeliegenden Corpus bezieht; eine weitergehende Beurteilung der generativen Adäquatheit ist dann zunächst ein Problem der Beurteilung und Optimierung der kognitiven Adäquatheit der Syntax.

Unsere Syntax ist nicht hinsichtlich jedes musikalischen Strukturaspekts analytisch adäquat, sobald man Lieder zu analysieren versucht, die nicht im Corpus enthalten sind. Die Inadäquatheit erstreckt sich unserer Erfahrung nach aber nicht auf die Periodenstruktur und die Transformationsregeln, sondern ausschließlich auf einige harmonische Details, die aufgrund zu restriktiver Kontextangaben nicht erfaßt werden können. Für Musikwissenschaftler als den kompetenteren Fachleuten liegt hier noch ein weites, erst in Ansätzen erschlossenes Arbeitsgebiet.

Kognitive Adäquatheit der Syntax

Eine musikalische Syntax, die wesentliche Eigenschaften besitzt, die zu Hypothesen führen, die nicht mit bereits vorliegenden Befunden und Modellen über die Struktur der Repräsentation und die Verarbeitungsprozesse beim Hören von Musik oder von einfachen Tonfolgen in Einklang stehen, besitzt geringe Chancen, sich als kogni-

tiv adäquat zu erweisen bzw. brauchbare neue Hypothesen über die kognitive Verarbeitung beim Musikhören anzuregen. Wir werden deshalb die kognitive Adäquatheit unserer Syntax so weit überprüfen, wie dies beim gegenwärtigen Erkenntnisstand möglich ist.

An dieser Stelle sei noch einmal betont, daß wir die Beziehung zwischen Syntaxmodell und kognitivem Modell nicht im Sinne der älteren psycholinguistischen Korrelationshypothese sehen (z.B. Miller, 1962): Das Syntaxmodell ist weder ein Modell der kognitiven Repräsentation noch ein Modell der kognitiven Verarbeitung beim Hören von Musik; seine Realität ist die der notierten Musik und nicht die der kognitiv repräsentierten Musik; streng genommen ist die Syntax noch nicht einmal ein System von Hypothesen über die Struktur der kognitiven Repräsentation, sondern eine Form der Beschreibung, die sich in ein System von Hypothesen übersetzen läßt, in das statt musikwissenschaftlich-linguistischer Konstrukte kognitive Konstrukte eingehen, ohne daß zwischen beiden Typen von Konstrukten eine eindeutige Beziehung zu bestehen braucht.

Daß z.B. die syntaktische Generierung nicht das Durchlaufen einer Repräsentationshierarchie beschreiben kann, wird daran deutlich, daß der Vorgang der Generierung einer Struktur immer simultan zwei oder drei Einheiten einer Ebene erzeugt; die klangliche Realisation dieser Einheiten stellt aber eine zeitliche Sukzession dar, und die kognitive Verarbeitung muß daher eher an der klanglichen Realisation als am Vorgang der syntaktischen Generierung orientiert sein.

Eine Voraussetzung für kognitive Adäquatheit, deren Erfüllung gleichzeitig einen Beleg für die generative Adäquatheit der Syntax darstellt, ist das Unvermögen kompetenter Hörer, zwischen unbekannten echten Volksliedmelodien und Melodien zu differenzieren, die durch die Syntax generiert wurden. In einem Experiment (Stoffer, 1979) konnten wir zeigen, daß die Hörer zwar zwischen syntaktisch richtigen Volksliedern (ihnen unbekannten echten und generierten Liedern) und solchen diskriminieren können, die in unterschiedlichem Ausmaß von den Regeln der Syntax abweichen, nicht aber zwischen echten und generierten Liedern.

Der wesentliche Aspekt der Syntax, der zu psychologischen Hypothesen über die Wahrnehmung von Musik veranlaßt, ist die hierarchische Struktur der durch sie erzeugten oder mit ihrer Hilfe zu analysierenden Melodien. Diese strukturelle Ähnlichkeit mit natürlichen Sprachen veranlaßte bereits Lashley (1951) zu der Hypothese, daß die kognitive Repräsentation musikalischer Strukturen eine hierarchische sein müsse. Diese auch von Sundberg & Lindblom (1976) geäußerte Hypothese ist Gegenstand der weiteren Experimente unserer Untersuchung.

Diese Hypothese hat drei Aspekte: 1. kann man fragen, ob die Repräsentationsstruktur bei der Wahrnehmung bereits eine hierarchische ist, 2. kann man fragen, ob die Gedächtnisrepräsentation eine hierarchische ist, 3. kann man auch noch fragen, ob das bei musikalisch kompetenten Hörern vorhandene implizite Wissen über musikalisch-syntaktische Regelmäßigkeiten eine hierarchische Form der Organisation zeigt. Wir wollen in unseren Experimenten das Hauptgewicht auf den ersten Aspekt legen.

Die Hypothese einer hierarchischen Wahrnehmungsrepräsentation ist aufgrund unserer Experimente 1 und 2, die wesentliche Aspekte des Hierarchiemodells mit positivem Ergebnis überprüft haben, bereits als plausibel zu betrachten. Zwei Teilergebnisse der Experimente 1 und 2 sind es, die, wenn man sie im Zusammenhang betrachtet, die Hypothese einer hierarchischen Repräsentation wahrscheinlich machen. In Experiment 1 fanden wir, daß der nicht zu beachtende erste Ton gar nicht unbeachtet bleiben kann. Dieser Befund ist nur damit zu erklären, daß die Verarbeitung auf dem Niveau der Gesamttonfolge beginnt und erst in einem zweiten Schritt eine Fokussierung der aufgabenrelevanten Detailinformation stattfindet, d.h. es müssen in diesem Fall mindestens zwei Repräsentationsniveaus vorgelegen haben, die sich hinsichtlich der Korngröße der repräsentierten Information unterscheiden.

In Experiment 2 fanden wir darüber hinaus, daß die Identifikation der Gesamttonfolge unabhängig von der Komplexität der Intervallkonfiguration einer Tonfolge weniger Zeit in Anspruch nimmt als die Identifikation der Intervallrichtung eines der beiden Tonpaare, obwohl ein einzelnes Intervall aus weniger Elementen besteht als die Gesamttonfolge. Dieses Ergebnis spricht eindeutig gegen eine Verarbeitung, die bei den Elementen beginnt und eine Repräsentation des Wahrgenommenen durch Integration der Elemente aufbaut. Auch hier zeigt sich also wieder, daß offenbar die Repräsentation hierarchisch von oben nach unten durchlaufen wird. Die Hypothese der hierarchischen Repräsentation sollte sich aber mit solchem Reizmaterial noch besser plausibel machen lassen, dessen hierarchische Struktur differenzierter ist, als die derart kurzer Tonfolgen, wie wir sie in den beiden ersten Experimenten verwendet haben.

Ein weiteres Argument für die kognitive Adäquatheit der Syntax besteht in der Ähnlichkeit unseres Transformationskonzepts mit dem der Invarianz. Die Transformationen "Wiederholung", "Sequenz" und "Umkehrung" konstituieren bei akustischer Realisation und Wahrnehmung durch einen Hörer Invarianzen höherer Ordnung bezüglich der Intervallkonfigurationen. Diese Transformationen spezifizieren die Informationen, die als Invarianzen direkt wahrnehmbar sein sollten:

"The perception of the persistent identity of an object that under goes a change over space-time logically presupposes that there is perceptual information available which specifies the continuity of the transformation underlying the change. ... Thus, it is the perceptual information for the continuity of the transformation that is of paramount importance in accounting for the perception of events rather than the perceived identity of the structure involved in the event transformation" (Shaw & Pittenger, 1978).

Wenn also, wie Shaw & Pittenger (1978) annehmen, die Transformationen es sind, die als Kontinuität der Veränderung eines sich in der Zeit entwickelnden Ereignisses wahrgenommen werden, dann muß die Repräsentationsstruktur zumindest solcher Tonfolgen, die auf mehreren Strukturniveaus Transformationen enthalten, eine hierarchische sein.

Ein weiteres Argument, das ebenfalls für eine hierarchische Repräsentation bei der Wahrnehmung von Musik spricht, sind die Befunde von Perkins (1974) und Sturges & Martin (1974) zur Rhythmuswahrnehmung, die sich am besten mit der Annahme einer hierarchischen Repräsentation rhythmischer Strukturen erklären lassen. Auch die Befunde von Deutsch (1980b) und Jones et al. (1978) zum Lernen hier-

archisch nach dem Merkmal "Tonhöhe" organisierter Tonfolgen sprechen indirekt für die Annahme einer hierarchischen Repräsentation des Wahrgenommenen. Deutsch (1980b) und Jones et al. (1978) wiesen zwar "nur" nach, daß die Gedächtnisrepräsentation eine hierarchische sein muß, aber wie soll diese entstehen, wenn nicht bereits aufgrund einer hierarchisch strukturierten Repräsentation der wahrgenommenen Tonfolgen? Die Autorinnen nehmen zwar an, daß die Struktur der Gedächtnisrepräsentation erst in ihrer hierarchischen Form aufgebaut werden müsse, aber dieser Aufbau muß ja gesteuert verlaufen, also in Abhängigkeit von den wahrgenommenen Strukturen und/oder in Abhängigkeit von kognitiven Strukturen, wie Schemata oder Frames. Letztere aber können nur dann selbst hierarchisch strukturiert sein und damit eine hierarchische Gedächtnisrepräsentation aufbauen, wenn sie auf der Grundlage der Wahrnehmung hierarchischer Strukturen einmal entstanden. Letztlich kann es also nur die Struktur der Wahrnehmungsrepräsentation sein, die für alle anderen Formen kognitiver Strukturen dafür entscheidend ist, welche Organisation die Strukturen zeigen.

Sowohl einige experimentelle Befunde als auch theoretische Argumente lassen die Hypothese, die Repräsentationsstruktur bei der Wahrnehmung von Musik sei eine hierarchische, plausibel erscheinen. Da diese Hypothese mit der aus der Syntax abzuleitenden Hypothese übereinstimmt, kann die Syntax hinsichtlich dieses Aspekts vorläufig als kognitiv adäquat bezeichnet werden.

1.4 ZUSAMMENFASSUNG

Gibson's (1966a) Strategie folgend, die Strukturen der Wahrnehmungs-

welt zuerst zu erforschen, bevor man Hypothesen über die Strukturen ihrer Repräsentation und über kognitive Prozesse aufstellt, haben wir eine generative Transformationssyntax für einen musikalischen Formtyp formuliert, die die Strukturen eines begrenzten Corpus deutscher Volks- und Kinderlieder als hierarchisch gegliederte musikalische Periode beschreibt.

Die Strukturen der syntaktischen Beschreibung dieser Lieder legen die Hypothese nahe, daß die Struktur der Repräsentation wahrgenommener Volksliedmelodien hierarchisch sein sollte. Die kognitive Adäquatheit der Syntax hinsichtlich der Struktur der von ihr zu erzeugenden Melodien und die psychologische Plausibilität der Hypothese einer hierarchischen Repräsentation von Musik wurden auf dem Hintergrund des Hierarchiemodells und einiger Befunde aus den Experimenten 1 und 2 sowie anhand von Befunden zur hierarchischen Repräsentation einfacher Tonfolgen aus der Literatur diskutiert. Danach kann die Hypothese einer hierarchischen Repräsentation bei der Wahrnehmung von Musik als sehr wahrscheinlich betrachtet werden. Die im folgenden zu berichtenden Experimente 3 bis 5 sind der Überprüfung dieser Hypothese gewidmet.

2. EXPERIMENTE ZUR REPRÄSENTATIONSSTRUKTUR BEIM HÖREN VON MELODIEN

In Experiment 3 übertragen wir die aus der Psycholinguistik stammende Methode der Lokalisation eines einem Satz überlagerten Klickgeräuschs auf die Untersuchung von Wahrnehmungsprozessen beim Hören von Melodien. Gezeigt werden soll in diesem Experiment, daß die Richtung und Stärke der Fehllokalisation eines einer Melodie überlagerten Klickgeräuschs systematisch in Abhängigkeit von der Position des Klicks variiert. Die Systematik der Variation ist aufgrund der syntaktischen Beschreibung der Melodien vorherzusagen. Ein Vergleich dieser Vorhersagen mit den empirischen Fehllokalisationsdaten gibt Aufschluß über die kognitive Adäquatheit der syntaktischen Beschreibung, und die Verteilung der Fehllokalisationen läßt Schlüsse auf die Struktur der kognitiven Repräsentation von Melodien beim Hören zu.

Experiment 4 thematisiert die Frage, wie sich die Repräsentation der Melodiestrukturen zwischen musikalisch-syntaktisch trainierten Hörern und musikalischen Laien unterscheidet. Die Methode ist hier ähnlich der des zweiten Experiments; statt eines Lokalisationsurteils über die wahrgenommenen Klickpositionen wird in Experiment 4 als abhängige Variable die Reaktionszeit (RT) für das Entdecken eines Klicks erhoben.

Mit Experiment 5 beschreiten wir schließlich einen methodisch anderen Weg. Hier werden von den Hörern same/different-Vergleiche zwischen Ausschnitten verschiedener Größe vorgenommen, wobei die zeitliche und die syntaktische Nähe der zu vergleichenden Aus-

schnitte variiert wird. Wenn die auf der Grundlage der Syntax abzuleitenden Hypothesen über die Struktur der Repräsentation beim Hören von Melodien zutreffen, dann sollte ein Vergleich zwischen syntaktisch definierten Ausschnitten unabhängig von ihrer zeitlichen Nähe eine um so längere Verarbeitungszeit benötigen, je geringer ihre syntaktische Nähe ist.

Um solche Experimente durchführen zu können, bedarf es einer Stichprobe von Vpn, die die syntaktischen Strukturen der ihnen darzubietenden Melodien genauestens kennen. Diesen Versuchen ging deshalb ein mehrwöchiges Training voraus, in dessen Verlauf die Vpn der experimentellen Stichprobe ein differenziertes syntaktisches Wissen über die Strukturen solcher Melodien erwarben, die ihnen in den anschließenden Experimenten dargeboten werden sollten.

2.1 SYNTAXTRAINING

Mit der Durchführung eines Syntaxtrainings verfolgten wir die folgenden Ziele: 1. Die Vpn der Experimente 3 bis 5 sollten ein differenziertes musikalisch-syntaktisches Wissen erwerben, um zum Zeitpunkt der Durchführung der Experimente über geeignete kognitive Strukturen zu verfügen, die es ihnen gestatten, relativ schwierige Strukturerkennungsaufgaben mit einem hohen Maß an Sicherheit und Schnelligkeit zu bewältigen, was insbesondere für die Erhebung von RT-Daten eine notwendige Voraussetzung ist. 2. Um die Varianz in den zu erhebenden abhängigen Variablen möglichst gering zu halten, die auf Unterschiede in der musikalischen Kompetenz der Vpn zurückzuführen sein könnte, ist in dieser Hinsicht eine relativ homogene Stichprobe notwendig. Da Musikalitätstests notorisch

unreliabel und meist von unbekannter Validität sind (vgl. de la Motte-Haber, 1972; Shuter, 1968), benötigen wir ein Selektionskriterium, dessen Validität der Messung derjenigen Fähigkeiten, die in den Experimenten von den Vpn erwartet werden, möglichst hoch ist. Ein solches Selektionskriterium können wir nur dadurch gewinnen, daß wir versuchen, einer Stichprobe von Vpn, die bereits einige musikalische Grundkenntnisse mitbringen, diese Fähigkeiten im Rahmen eines Trainings zu vermitteln, anschließend den Erfolg des Trainings messen und aufgrund dieser Messung eine Homogenisierung der Stichprobe vornehmen. 3. Um mit den Vpn über Instruktionen und Aufgabenbeschreibungen kommunizieren zu können, so daß den Vpn verbal mitgeteilt werden kann, auf welche Strukturaspekte und Ausschnitte einer Melodie sie sich konzentrieren sollen, ist es erforderlich, Begriffe zur Beschreibung musikalischer Strukturen zu definieren und das Verständnis der Vpn bei ihrer Anwendung zu trainieren und zu überprüfen. Ein Teil des Trainings bestand aus diesem Grunde darin, musikalische Strukturbegriffe zu definieren und ihre Anwendung beim Erkennen struktureller Beziehungen in Melodien zu üben.

Da die vollständige Syntax der Volks- und Kinderlieder viel zu komplex wäre, um sie Vpn in einem relativ kurzen Übungsprogramm im Detail beibringen zu können, beschränken wir uns auf eine Teilmenge der syntaktischen Regeln, die so etwas wie prototypische Liedstrukturen erzeugen können. Alle ausgelassenen Regeln sind in der Syntax als optionale Erweiterungen oder Kürzungen der prototypischen Periodenstruktur gekennzeichnet. Nur die optionalen Erweiterungen durch Anschlußgliedwiederholung und Motivsequenz

(Regeln 4 und 9, jeweils 2. Zeile, vgl. Anh. 1) wurden in das Trainingsprogramm aufgenommen, weil hierbei das Erkennen der Transformationen "Wiederholung" und "Sequenz" unter erschwerten Bedingungen (Abweichung von der typischen achttaktigen Periode) geübt werden kann.

Um die uns im gegenwärtigen thematischen Zusammenhang weniger interessierenden rhythmusabhängigen Beeinflussungen der Wahrnehmungsprozesse so gut wie möglich konstant zu halten, verwandten wir im Training (und auch in den anschließenden Experimenten) nach einer Demonstration einiger struktureller Eigenschaften an bekannten Volksliedern ausschließlich Melodien im 4/4-Takt ohne Auftakt und mit gleichlangen Notenwerten (vier Viertelnoten). Die dadurch entstehende Einbuße an ästhetischem Reiz der Melodien mußte leider in Kauf genommen werden.

Beschreibung des Trainingsprogramms

Das Trainingsprogramm war so aufgebaut, daß die Periodenstruktur von der obersten (Periode) zur untersten Hierarchieebene (Viertelnoten, Betonungsränge) durchlaufen wurde und für jeden Übergang von einer zur anderen Ebene die Regeln angegeben wurden. Gleichzeitig wurden die Transformationen, die auf einer jeden Ebene horizontale Beziehungen zwischen den strukturellen Einheiten stiften, angegeben.

Alle Regeln wurden zunächst dadurch eingeführt, daß den Vpn die Aufgabe gestellt wurde, eine durch eine Regel spezifizierte Be-

ziehung in einem vorgespielten und im Notentext zu verfolgenden Beispiel zu entdecken. Dann wurde die gefundene Beziehung in Form einer Regel formuliert und das Erkennen dieser Beziehung geübt, wobei zuerst noch die vorgespielten Melodien im Notentext zu verfolgen waren, später aber mußten sich die Vpn ganz auf ihr Gehör verlassen. Der vollständige Trainingstext mit den Notenbeispielen sowie das Arbeitsheft für die Vpn mit Notenbeispielen ist im Anh. 2 zusammengestellt.

Durchführung des Trainings

Das Training wurde in vier dreistündigen Sitzungen (mit einer halbstündigen Pause in der Mitte einer jeden Sitzung) in wöchentlichem Abstand durchgeführt. Jeweils vier bis sechs Vpn nahmen gemeinsam an einer Sitzung teil. Die mit dem Klavier auf Tonband aufgezeichneten Melodien wurden über Lautsprecher dargeboten.

Das Training besaß die Form einer programmierten Unterweisung. Den Vpn stand ein Lehrtext und ein Arbeitsheft für ihre Antworten zur Verfügung (vgl. Anh. 2). Die Melodiebeispiele wurden vom Vl vorgespielt, wenn alle Vpn die entsprechende Textstelle im Lehrtext erreicht hatten. Der Vl [1] besprach auch die von den

---

[1] Vl war Herr Rudolf Vogelhuber, der auch die Tonbandaufnahmen der Melodiebeispiele mit dem Klavier besorgte. Er erwies sich als musikalisch sehr versierter Vl, der alle Fragen der Vpn kompetent und verständlich zu beantworten wußte.

Vpn gegebenen Aufgabenlösungen mit ihnen und gab, falls notwendig, Erläuterungen über den Lehrtext hinaus. Am Ende einer jeden Trainingssitzung wurde ein kleiner Test durchgeführt, der 1. weitere Übungsmöglichkeiten bot und 2. eine Überprüfung des Verständnisses ermöglichte. Der Vl besprach mit den Vpn anschließend die Testfragen und gab Rückmeldung über die Richtigkeit der Antworten, wobei er die entsprechenden Melodien ein zweites Mal vorspielte.

In der letzten Trainingssitzung wurde ein längerer Test mit 54 Übungsmelodien und Strukturerkennungsaufgaben in Form eines Multiple-Choice-Tests durchgeführt. Von den 48 Vpn, die das Training beendeten, qualifizierten sich 44 Vpn für die Teilnahme an den weiteren Experimenten durch mehr als 85 % richtiger Antworten. Bei den Vpn handelte es sich um Studenten verschiedener Fachrichtungen, die teilweise für ihre Teilnahme mit 8 DM pro Stunde bezahlt wurden; die Psychologiestudenten erhielten Testate für abgeleistete Vpn-Stunden. Alle Vpn spielten seit mehr als drei Jahren mindestens ein Musikinstrument.

2.2 EXPERIMENT 3: LOKALISATION EINES EINER MELODIE ÜBERLAGERTEN KLICKS

2.2.1 EINLEITUNG

Die Hypothese einer hierarchischen Struktur von Repräsentationen beim Hören von Melodien wollen wir zuerst mit der experimentellen Methode überprüfen, die in den ersten Untersuchungen zur "psychologischen Realität der Grammatik" angewandt wurde, nämlich mit der Klicklo-

kalisationsmethode. Diese erstmals von Lodefoged & Broadbent (1960) beschriebene Methode besteht darin, einer Folge akustischer Signale an einer Stelle ein kurzes Klickgeräusch zu überlagern, das vom Hörer zu entdecken und in bezug auf die Reihenfolge der Signale zu lokalisieren ist. Dem Rationale dieses experimentellen Paradigmas liegt die theoretisch auf recht unterschiedliche Weise zu begründende Annahme zugrunde, das akustische Reizmuster werde nicht zeitanalog kontinuierlich, sondern in physisch durch Pausen und Betonungen und/oder strukturell abgegrenzten Segmenten verarbeitet, wobei der Klick von dieser Verarbeitung solange ausgeschlossen wird, bis die Kodierung eines Segments abgeschlossen ist. Dadurch geht phänomenal die zeitliche Beziehung des Klicks zu den Elementen innerhalb eines Segments verloren, so daß der Klick nur davor oder dahinter lokalisiert werden kann. Positionen, an denen ein Klick besonders exakt im Vergleich zu anderen Positionen des Reizmusters zu lokalisieren ist, kennzeichnen somit Segmentgrenzen. Richtung und Betrag der eventuell auftretenden Fehllokalisationen weisen auf die der Klickposition am nächsten liegende Segmentgrenze hin.

Dieses experimentelle Paradigma wurde in psycholinguistischen Experimenten dazu verwandt, die Größe und interne Gliederung von Einheiten der Sprachwahrnehmung zu bestimmen und diese identifizierten Einheiten mit den syntaktisch definierten in Beziehung zu setzen. Wir wollen zunächst die wichtigsten Untersuchungen, die mit dieser Methode gearbeitet haben, darstellen, wobei wir besonderes Gewicht auf die Frage der Validität dieser Methode zur Bestimmung von Einheiten der auditiven Wahrnehmung legen und

einige methodische Probleme berücksichtigen, die mit der Frage der Validität der Methode zusammenhängen.

Die Anwendung des Klickparadigmas in der psycholinguistischen Forschung

Ladefoged & Broadbent (196o) beabsichtigten nachzuweisen, daß die Wahrnehmung gesprochener wie die geschriebener Sprache nicht kontinuierlich erfolgt, sondern in Einheiten, die größer sind als einzelne Sprachlaute. Sie postulierten einen zentralen Mechanismus, der ein "sampling" von Sprachlauten vornimmt, d.h. einen Mechanismus, der Gruppen von Sprachlauten bildet und jeweils eine Gruppe als Ganze weiterverarbeitet. Das sollte dazu führen, daß die physische Reihenfolge der Laute nicht mit der wahrgenommenen Reihenfolge übereinzustimmen braucht, wenn durch das "sampling" die Sukzessivität der Laute aufgehoben wird und sie simultan verarbeitet werden.

Der theoretische Ansatz basiert auf Titchener's "law of prior entry", wonach ein Reiz, auf den man eingestellt ist, weniger Zeit benötigt, um einen bewußten Einfluß auszuüben, als ein ähnlicher Reiz, auf den man nicht vorbereitet ist (Titchener, 19o9). Dieses "Gesetz" bezieht sich allerdings auf Experimente, in denen zwei Stimuli über verschiedene Sinnesmodalitäten aufgenommen wurden. Wenn das "law of prior entry" aber auch für Reize gilt, die über dieselbe Modalität aufgenommen werden, dann - so die Folgerung der Autoren - sei dies ein Zeichen dafür, daß die Reize nicht sukzessiv verarbeitet werden, sondern ein "sampling" statt-

findet, das Gruppen simultan zu verarbeitender Reize herstellt, von denen der Klick ausgeschlossen wird.

Die Autoren führten Experimente durch, in denen einer Zahlenreihe oder einem Satz ein Klick überlagert war. Die Vpn sollten die Stelle angeben, wo der Klick zu hören war, indem sie das Wort oder die Zahl aussprachen, wo sie den Klick lokalisieren konnten bzw. die beiden Wörter oder Zahlen, zwischen denen sie den Klick hörten. In ihrem ersten Experiment verglichen sie zwei Bedingungen: In der ersten war den Vpn der Text unbekannt, während die Vpn in der zweiten Bedingung den Text vorher kennengelernt hatten, bevor sie ihn satzweise vom Tonband hörten. Für die Fehllokalisation des Klicks wurde pro Satz ein Score gebildet, der die Zahl der Einheiten angab, um den der Klick zu früh oder zu spät lokalisiert wurde. Jedes Wort und jeder Wortzwischenraum galten hierbei als Einheit. Für die Gruppe, die den Text vorher nicht kannte, wurde ein mittlerer Score pro Satz von 1.8 und für die Gruppe, die den Text vorher kannte, ein mittlerer Score von 2.3 ermittelt. Dieser Unterschied war nicht signifikant, während jeder Score signifikant von Null verschieden war. Diese Fehllokalisation von durchschnittlich zwei Einheiten zeigt, daß die Klicks zusammen mit dem Wort gehört wurden, das dem Wort vorausging, mit dem der Klick tatsächlich gemeinsam auftrat.

Die Befunde von Ladefoged & Broadbent (196o) zeigen, daß eine Tendenz besteht, ein einer auditiv dargebotenen, sinnvollen Lautfolge überlagertes Signal f r ü h e r zu lokalisieren als es seinem tatsächlichen Auftreten in der Lautfolge entspricht, und daß dieser Effekt unabhängig von der vorherigen Kenntnis der Lautfolge

und von der Art des überlagernden Signals ist. Die Autoren interpretieren diese Befunde als Beleg dafür, daß die Verarbeitung bei der Wahrnehmung von Sprache von diskreten Einheiten ausgeht, die größer als einzelne Laute sind.

Mit demselben experimentellen Paradigma untersuchten auch Fodor & Bever (1965) die Frage nach der "psychischen Realität linguistischer Segmente", indem sie nach der Entsprechung der Verarbeitungssegmente mit den Segmenten einer linguistischen Sprachanalyse fragten. Fodor & Bever (1965) stellten die Hypothese auf, daß die Einheiten bei der Sprachwahrnehmung den Konstituenten eines Satzes entsprechen. Sie analysierten die Satzstrukturen der von Ladefoged & Broadbent (196o) verwandten Sätze in Beziehung zur Fehllokalisation der Klicks und fanden dort ihre Hypothese bestätigt, denn die Hauptkonstituentengrenzen, also die Grenzen zwischen Nominal- und Verbalphrasen, lagen bei Ladefoged & Broadbent in den meisten Fällen vor dem objektiven Auftreten des Klicks. Sie führten nun Experimente durch, in denen ein Klick bis zu drei Silben vor oder hinter der Hauptkonstituentengrenze eines Satzes auftreten konnte oder genau mit der Konstituentengrenze zusammenfiel. Die Vpn sollten sich einen jeden Satz genau anhören, ihn niederschreiben und dann die Stelle kennzeichnen, an der sie den Klick gehört hatten.

Die Befunde zeigten, daß die Klicks an der jeweils nächsten Konstituentengrenze lokalisiert wurden, obwohl sie objektiv davor oder dahinter auftraten. Die Häufigkeit richtiger Lokalisationen war bedeutend größer als die der Fehllokalisationen, wenn der Klick an der Hauptkonstituentengrenze auftrat. Die Verteilung akustischer

Pausen in den Sätzen zeigte keinen Zusammenhang mit den Häufigkeiten der Fehllokalisationen. Fodor & Bever (1965) zogen aus diesen Experimenten den Schluß, daß die syntaktischen Einheiten mit den Einheiten der Sprachwahrnehmung übereinstimmen und daß die Klickverschiebung dem Erhalt der Integrität dieser Einheiten dient.

Garrett, Bever & Fodor (1965) wiederholten die Experimente von Fodor & Bever (1965) mit der Absicht auszuschließen, daß solche Faktoren wie Intonation, Rhythmus und Betonung für die Klickverschiebung von Bedeutung sein könnten. Dabei verwandten sie mehrdeutige Satzteile, deren Bedeutung erst durch den Kontext deutlich wird. Ein mehrdeutiger Satzteil wurde auf Band aufgenommen und mit verschiedenen sprachlichen Kontexten zusammengeschnitten. Garrett et al. (1965) replizierten die Befunde von Fodor & Bever, indem sie nachweisen konnten, daß die Fehllokalisation alleine von der Konstituentenstruktur eines Satzes und nicht von der Intonation, vom Rhythmus oder von der Betonung abhängig ist.

Im Anschluß an diese ersten Untersuchungen zu den Einheiten der Sprachwahrnehmung entspann sich eine Diskussion darüber, ob es sich bei diesen Einheiten um Oberflächen- oder Tiefenstruktursegmente handeln könnte. Bever, Lackner & Kirk (1969) erhielten Befunde, die für Tiefenstruktursegmente sprachen, Chopin, Smith & Abrahamson (1972) sowie Toppino (1974) erhielten Befunde, die mehr für die Bedeutung der Oberflächenstruktur sprachen. Diese Frage ist zu spezifisch für Sprache, als daß sie uns hier näher beschäftigen sollte. Interessanter für unsere Absicht, das Klick-

Paradigma für die Untersuchung der Repräsentationsstruktur beim Musikhören einzusetzen, sind methodenkritische Untersuchungen und solche zur funktionalen Erklärung des Fehllokalisationseffekts.

Eine Alternativerklärung für die berichteten Befunde wäre die, daß es sich hierbei um einen Gedächtniseffekt handelt, da die Vpn ja die Sätze kurzfristig behalten müssen, um sie niederschreiben zu können. Wenn sie erst im Anschluß daran die Position des Klicks markieren, könnten sie ihn deshalb falsch lokalisiert haben, weil sie seine Position inzwischen vergessen haben.

Diese Alternativerklärung wird durch die Befunde von Ladefoged (1967) nahegelegt, der seinen Vpn sagte, die Sätze enthielten unterschwellige Klicks, aber in Wirklichkeit waren überhaupt keine Klicks vorhanden. Auch in dieser Situation werden die angeblichen Klicks an den Hauptkonstituentengrenzen lokalisiert.

Dieser Befund rief nun eine Anzahl methodenkritischer Experimente auf den Plan, deren Hauptziel darin bestand, nachzuweisen, daß der Fehllokalisationseffekt nicht nur durch systematische Reaktionstendenzen der Vpn entstanden sein kann. Das wichtigste Experiment mit dieser Zielsetzung ist das von Bever (1973).

In diesem Experiment wurden den Vpn an einem Ohr Sätze und am andern Ohr Klicks in weißem Rauschen dargeboten. In einigen Fällen wurde der Klick fortgelassen und nur weißes Rauschen dargeboten, um die Bedingung von Ladefoged (1967) mit den anderen

vergleichen zu können. Der postulierte Effekt aufgrund von systematischen Reaktionstendenzen ließ sich damit durch Vergleich der Bedingung ohne Klicks mit denen mit Klicks ermitteln. Wenn Klicks auftraten, erschienen diese eine Silbe vor oder hinter der Hauptkonstituentengrenze oder genau in ihr.

Es wurden drei Antwortbedingungen mit den beiden Klick-Bedingungen kombiniert: 1. Normale Klick-Bedingung: Die Vpn hörten einen Satz und schrieben ihn anschließend nieder, wobei sie die Position des Klicks markierten. 2. Den Vpn wurde ein Transskript der Sätze gegeben, das sie während des Hörens verfolgen und dabei die Klickposition markieren konnten. 3. Die Vpn wußten nicht im voraus, ob sie den Satz niederschreiben mußten oder ein Transskript bekamen, weshalb sie immer den Satz aufmerksam hören mußten; in einigen Versuchsdurchgängen erhielten sie nach dem Hören des Satzes das Transskript, das den Bereich der Position des Klicks auf $\pm$ 2 Silben genau angab.

Die Befunde zeigen, daß in der ersten Antwortbedingung eine signifikante Differenz in der Verteilung der Klicklokalisationen zwischen den beiden Klick-Bedingungen auftrat, und zwar war die Häufigkeit von Fehllokalisationen in Richtung der Hauptkonstituentengrenzen größer, wenn Klicks wirklich dargeboten wurden als wenn keine Klicks auftraten. Diese Asymmetrie der Verteilung über die beiden Klick-Bedingungen trat aber nicht bei der zweiten Antwortbedingung auf. D.h. es gibt einen Fehllokalisationseffekt unabhängig von einer offenbar auch vorhandenen systematischen Reaktionstendenz, aber nur dann, wenn die Vpn aufmerksam dekodieren müssen, wie in Bedingung 1. Diesen Befund erhielten auch Bertelson & Tisseyre (197o).

wenn der Fehllokalisationseffekt nur dann auftreten würde, wenn die Vpn den Klick in der Nähe der Konstituentengrenzen erwarten, sollte die Häufigkeit der Fehllokalisation genau zwischen den Konstituentengrenzen bei alleiniger Wirkung einer systematischen Reaktionstendenz für beide Klick-Bedingungen gleich sein, wenn eine ungefähre Information über die zu erwartende Position des Klicks gegeben wird, wie in Bedingung 3. Tatsächlich trat die gleiche Asymmetrie in der Häufigkeitsverteilung der Fehllokalisationen bei Vergleich der beiden Klick-Bedingungen innerhalb der Antwortbedingung 3 auf. Diese Befunde bestätigen, daß der Effekt tatsächlich durch die syntaktische Segmentierung entsteht und unabhängig von einer möglicherweise zusätzlich wirksamen systematischen Reaktionstendenz zu sein scheint.

Daß das aufmerksame Beachten der Sprache eine notwendige Bedingung für das Auftreten eines sich an den Konstituentengrenzen orientierenden Fehllokalisationseffekts ist, wurde unabhängig von Bever (1973) ungefähr zur gleichen Zeit von Seitz & Weber (1974) bestätigt.

Eine weitere Bestätigung dafür, daß eine systematische Reaktionstendenz alleine den Effekt nicht erklären kann, wurde in jüngster Zeit von Dalrymple-Alford (1976) gegeben. Der Grundgedanke für das Experiment war der folgende: Systematische Reaktionstendenzen sollten sich umso stärker bemerkbar machen, je weniger sicher die Vpn glauben, den Klick lokalisieren zu können, und die Stärke der Fehllokalisation sollte bei Wirkung einer systematischen Reaktionstendenz dann am stärksten sein, wenn die Urteilsunsicherheit am größten ist. Die Vpn hatten die Aufgabe, den Klick im Satz zu lo-

kalisieren, indem sie sich zuerst den Satz anhörten und dann auf einem ihnen nach dem Hören präsentierten Transskript die Position markierten. Wenn sie über die Position des Klicks unsicher waren, konnten sie mehrere Markierungen vornehmen. Die Anzahl der pro Satz vorgenommenen Markierungen wurde als Grad der Urteilsunsicherheit gewertet, aber nur die erste Markierung wurde als Lokalisationsurteil gewertet.

Wenn alleine eine systematische Reaktionstendenz vorliegen sollte, dann sollte die Häufigkeit von Fehllokalisationen in Richtung auf die Konstituentengrenzen bei unsicheren Urteilen größer als bei sicheren sein. Aber genau das umgekehrte Ergebnis stellte sich ein: Eine systematische Reaktionstendenz kann also nicht die Basis für den Fehllokalisationseffekt darstellen.

Abrams & Bever (1969) wiesen nach, daß der Fehllokalisationseffekt nicht auf das Vergessen der Klickposition zurückzuführen ist. Im Klickversuch reagiert die Vp normalerweise erst kurze Zeit nach dem Hören des ganzen Satzes. Diese Bedingung erlaubt nicht eine Identifikation der Verarbeitungsstufe, die für die Fehllokalisation der Klicks verantwortlich sein kann, weil die Reaktion erst nach einer vollständigen Verarbeitung des gesamten Inputs abgegeben wird. Aus diesem Grund verwandten die Autoren eine Form der Reaktion auf den Klick, die von eventuellen Gedächtniseffekten nicht beeinflußt wird, nämlich die Reaktionszeit für die Detektion des Klicks.

Der Versuchsaufbau war der übliche, lediglich mit dem Unterschied, daß die Vpn vor dem Niederschreiben des Satzes durch Betätigen

einer Recktionstaste so schnell wie möglich die Detektion des Klicks anzeigen mußten. Außerdem wurde im Unterschied zu ähnlichen Experimenten jeder Satz dreimal dargeboten.

Ladefoged & Broadbent (196o) hatten angenommen, daß die Segmentgrenzen diejenigen Stellen sind, an denen der Hörer in der Lage ist, dem Satz ein Minimum an Aufmerksamkeit zu schenken, so daß er anderen Stimuli mehr Aufmerksamkeit zuwenden kann, als bei der Verarbeitung eines Segments. Wenn dies der Fall wäre, dann sollte die Reaktionszeit auf einen Klick zwischen Konstituentengrenzen kürzer als innerhalb eines Konstituenten sein.

Diese Hypothese konnten Abrams & Bever nicht so eindeutig belegen, denn die Reaktionszeiten für den ersten Darbietungsblock von Sätzen waren für Klicks vor den Konstituentengrenzen am längsten und hinter den Konstituentengrenzen am kürzesten. Erst vom zweiten Darbietungsblock an war die kürzeste Reaktionszeit die auf den Klick zwischen Konstituentengrenzen, aber der Unterschied zu den anderen Klickpositionen war statistisch nicht bedeutsam. Auch bei Eliminierung der unterschiedlichen Betonungen der Wörter eines Satzes konnte die Hypothese nicht bestätigt werden.

Die Autoren interpretierten ihre Befunde trotzdem als Beleg dafür, daß die syntaktische Struktur eines Satzes Verarbeitungseinheiten festlegt, indem sie annahmen, daß während des Hörens einer Verarbeitungseinheit lediglich Information akkumuliert wird und die Kodierung einer Einheit erst am Ende stattfindet. Aus diesem Grunde ist die Kapazität des Kodierungsmechanismus

am Anfang einer Einheit weniger stark ausgelastet als am Ende, so daß die Lokalisation eines Klicks am Anfang einer Einheit schneller als an ihrem Ende erfolgen müßte.

Als Erklärungskonstrukt wurde also von Abrams & Bever eine sich über die Länge eines sprachlichen Segments erstreckende Zunahme der Kapazitätsauslastung verwendet, die gerade dann am höchsten sein soll, wenn eine Segmentgrenze erreicht wird, so daß an der Stelle die Kapazitätsbeanspruchung am höchsten ist, wenn die akustische Information über das nachfolgende Segment gerade akkumuliert wird.

Von einer ähnlichen theoretischen Vorstellung gehen Holmes & Foster (197o) aus. Auch sie versuchten, durch Reaktionszeitmessung für die Detektion eines dem Satz überlagerten Klicks die Hypothese zu testen, daß es sich bei diesem Effekt um einen Wahrnehmungseffekt und nicht ein auf Gedächtnisprozesse zurückführbares Phänomen handelt. Im Gegensatz zu Abrams & Bever (1969) erhielten Holmes & Foster (197o) Befunde, die mit dieser Hypothese im Einklang standen. Die Detektionszeiten für Klicks zwischen den Hauptkonstituentengrenzen waren signifikant kürzer als die für Klicks davor oder dahinter. Klicks, die zwischen kleinere Konstituentengrenzen fielen, wurden zwar geringfügig schneller erkannt als Klicks, die innerhalb eines Konstituenten lagen, aber die Differenz zu letzteren war nicht statistisch signifikant.

Diese Befunde bestätigen, daß die Fehllokalisation von Klicks funktional auf einen Wahrnehmungsprozeß zurückzuführen ist. Diese Interpretation wird auch noch dadurch unterstützt, daß sich keine

signifikante Korrelation zwischen Genauigkeit der Satzreproduktion und Reaktionszeit nachweisen ließ. Wie auch Abrams & Bever interpretieren Holmes & Foster dieses Ergebnis als Zeichen für eine größere Kapazitätsauslastung innerhalb eines Segments als zwischen Segmenten, aber aufgrund ihrer Befunde, die ja nicht mit denen von Abrams & Bever übereinstimmen, mußten sie annehmen, daß die Kapazitätsauslastung nur bis zur Segmentgrenze ansteigt und dann zwischen Segmentgrenzen wieder geringer wird.

In gleicher Weise konnten Holmes& Foster (1972) parallele Befunde aus Experimenten erklären, in denen die abhängige Variable nicht die Reaktionszeit für die Detektion eines Klicks, sondern die Genauigkeit der Klicklokalisation war.

Gegen die Interpretation, daß der Fehllokalisationseffekt die Segmentierung der Sprache durch den Hörer reflektiere, wendet sich Reber (Reber & Anderson, 197o; Reber, 1973a,b). Er führt alle Befunde auf systematische Reaktionstendenzen zurück. Wir können hier auf eine genauere Erörterung der Befunde von Reber verzichten, denn er verwandte eine Bedingung, unter der die systematischen Reaktionstendenzen besonders stark durchschlagen müssen: Die Vpn hörten einen Satz und brauchten nur anschließend auf einem Transskript, das sie während des Hörens allerdings nicht lesen konnten, die Klickposition zu markieren. Wie schon Bever (1973) und später Seitz & Weber (1974) gezeigt haben, ist der Effekt unter dieser Bedingung nur geringfügig durch die syntaktische Struktur der Sätze beeinflußt. Dies ist leicht dadurch erklärbar, daß die Vpn ja wissen, daß ein Klick

auftritt und sie außerdem wissen, daß sie auf die Sprache erst dann zu achten brauchen, wenn der Klick aufgetreten ist, um ihn in Beziehung zur Position im Satz setzen zu können. So können sie sich ganz auf die Detektion des Klicks konzentrieren, ohne allzu sehr durch die Verarbeitung des Satzes vor Auftreten des Klicks belastet zu sein.

Die methodische Diskussion um den Fehllokalisationseffekt hat für die abhängigen Variablen "Lokalisation des Klicks" und "Detektionszeit für Klicks" folgendes gezeigt:

1. Variable "Lokalisation des Klicks":

   a) Eine Komponente der Fehllokalisation von Klicks ist unabhängig von der syntaktischen Struktur des Satzes und offenbart eine systematische Tendenz, die zu einer zu frühen Lokalisation in Relation zur wahren Position führt.

   b) Diese systematische Tendenz setzt sich dann am stärksten durch, wenn die Versuchssituation es den Vpn gestattet, sich ganz auf die Lokalisation der Klicks zu konzentrieren und die Sprache erst dann zu beachten, wenn der Klick aufgetreten ist.

   c) Die Fehllokalisation eines Klicks orientiert sich erst dann an der syntaktischen Struktur des Satzes, wenn die Vpn die Aufgabe haben, den Satz vollständig zu dekodieren und ihn niederzuschreiben. Hierbei sind dann allerdings Gedächtnisfaktoren bei der Fehllokalisation nicht auszuschließen.

   d) Die Fehllokalisation reflektiert erst dann zweifelsfrei die zugrundeliegende systaktische Struktur, wenn die Vpn nicht vor dem Anhören des Satzes wissen, ob sie ihn sich merken müssen, um ihn niederzuschreiben, oder ob sie ein Transskript vorgelegt bekommen.

2. Variable "Detektionszeit für Klicks":
   Hier sind die Befunde weniger eindeutig, aber die Tendenz geht dahin, faßt man die Ergebnisse von Abrams & Bever (1969) sowie Holmes & Foster (197o) zusammen, daß die Reaktionszeiten eine Funktion der Position des Klicks in Relation zu den Hauptkonstituentengrenzen ist, wobei auf Klicks innerhalb eines Konstituenten langsamer reagiert wird als auf Klicks zwischen Konstituenten.

Übertragung des Klickparadigmas auf die Untersuchung der Repräsentationsstruktur beim Hören von Melodien

Gregory (1978) hat mit kurzen Tonfolgen ein Klickexperiment durchgeführt, in dem sechs Töne dargeboten wurden, die von einer Vpn-Gruppe als drei Paare von Tönen und von einer anderen Gruppe als zwei Dreiergruppen aufgefaßt werden sollten. Den Vpn wurde während des Anhörens der Tonfolgen das entsprechend gruppierte Notenbild vorgelegt, in dem sie auch die Klickposition zu markieren hatten. Eine Hälfte der Vpn hörte am rechten Ohr die Tonfolge und am linken den Klick, bei der anderen Hälfte wurden die Seiten vertauscht.

Die von Gregory (1978) gefundenen Klickverschiebungen waren minimal, gingen aber wie erwartet in Richtung der im Notenbild dargestellten Gruppengrenzen: Bei Auffassung der sechs Töne als zwei Dreiergruppen wurde die Mehrzahl der Klicks in Richtung der Grenze zwischen den Dreiergruppen verlagert, bei Auffassung als drei Zweiergruppen fand die Verlagerung in Richtung der nächsten Grenze zwischen zwei benachbarten Zweiergruppen statt. Interessant ist der

Befund von Gregory, daß der Klick im Durchschnitt s p ä t e r lokalisiert wurde, als seiner tatsächlichen Position entsprach, während in psycholinguistischen Klickexperimenten überwiegend eine v e r f r ü h t e Lokalisation gefunden wurde.

Für die beiden dichotischen Darbietungsbedingungen (Klick rechts, Tonfolge links oder umgekehrt) ergab sich der aus den psycholinguistischen Anwendungen des Klickparadigmas bekannte Asymmetrieeffekt: Bei Darbietung des Klicks am linken und der Tonfolge am rechten Ohr war der Fehllokalisationsbetrag deutlich niedriger als bei umgekehrter Darbietungsform. Dieser Befund spricht für die Hypothese der Hemisphärendominanz, wonach die linke Hemisphäre für zeitliche Analysen spezialisiert ist (Bever & Chiarello, 1974; Craig, 1979; Efron, 1963; Halperin, Nachshon & Carmon, 1973).

Daß die Klickverschiebungen in Gregory's Untersuchung nur minimal waren, ist angesichts der Befunde zur psycholinguistischen Anwendung des Klickparadigmas nicht erstaunlich, denn hier hat sich ja wiederholt gezeigt, daß Verschiebungen in Richtung der syntaktisch markierten Segmentgrenzen nur dann in bedeutendem Ausmaß auftraten, wenn die Vpn den Satz vollständig zu dekodieren hatten. In Gregory's Experiment dagegen brauchten die Vpn lediglich die Tonfolge in der Notendarstellung zu verfolgen und konnten sich ganz auf das Entdecken des Klicks konzentrieren. Die hier demonstrierte Klickverschiebung ist daher eher auf die visuelle Segmentierung des Notenbildes als auf eine auditive Einheitenbildung zurückzuführen. Ein Rückschluß auf die Struktur der Repräsentation der Tonfolgen ist aus diesem Grunde hier nicht möglich.

Ein Experiment, das diesen Rückschluß erlauben soll, muß folgendermaßen aufgebaut sein:

1. Die Hörer müssen sich in einer Doppelaufgabensituation befinden, in der die primäre Aufgabe ein Erkennen der Melodiestruktur verlangt und die sekundäre Aufgabe in der Lokalisation des Klicks besteht. Hierdurch kann in einer dichotischen Hörsituation gewährleistet werden, daß die Melodie konstant fokussiert wird, so daß die resultierenden Klickverschiebungen Charakteristika der Melodieverarbeitung und nicht primär der Verarbeitung des Klicks reflektieren.

2. Die Klicks dürfen nicht nur innerhalb, vor oder hinter der Hauptkonstituentengrenze auftreten, sondern auch innerhalb, vor und hinter allen anderen niedrigeren Konstituentengrenzen, damit eine eventuell vorhandene, mehrere Niveaus enthaltende Repräsentationsstruktur identifizierbar wird.

3. Die Konstituentengrenzen dürfen nicht durch irgendwelche sensorischen Hilfsmittel markiert werden, wie dies bei Gregory (1978) durch das Notenbild geschah, sondern müssen ausschließlich aus der Struktur der dargebotenen Melodien zu entnehmen sein.

Erklärung des Fehllokalisationseffekts im Rahmen des Hierarchiemodells

Zwei Effekte, die sich im Klickversuch zeigen lassen, je nach dem, ob man eine Detektions- oder eine Lokalisationsaufgabe stellt, sind

zu erklären: 1. die Abhängigkeit der Detektionszeit von der syntaktischen Struktur und 2. die Abhängigkeit des Fehllokalisationsbetrages von der syntaktischen Struktur. Da die Detektion des Klicks Voraussetzung für seine Lokalisation ist, können wir das Auffinden der Klickposition als zweistufigen Prozeß betrachten.

Die erste Stufe, die Detektion, besteht darin, die durch das Klickgeräusch produzierten Merkmalsveränderungen auf dem Hintergrund der durch den Melodieverlauf produzierten Merkmalsveränderungen zu entdecken. Betrachten wir zwei Ausschnitte eines beliebigen Repräsentationsniveaus und vergleichen die Dimensionalität der Merkmalsvariation innerhalb und zwischen diesen Ausschnitten, dann ist in dem Fall, in dem diese Ausschnitte strukturelle Binnenglieder der übergeordneten Repräsentationsebene sind (z.B. wenn es sich um zwei rhythmisch invariante Takte oder Takte mit einer transformationellen Invarianz handelt) die Merkmalsveränderung innerhalb der Ausschnitte dimensional komplexer als die zwischen Ausschnitten. Z.B. können innerhalb der Ausschnitte die Tonhöhe, die Intervallgröße und die Intervallrichtung variieren, aber zwischen Ausschnitten lediglich noch die Tonhöhe (z.B. bei Sequenzbeziehung zwischen Ausschnitten). Wenn die Detektion im Entdecken einer Merkmalsänderung besteht, dann müßte sich die durch den Klick produzierte Merkmalsänderung vom Hintergrund der Merkmalsänderungen i n n e r h a l b der Ausschnitte weniger deutlich abheben als auf dem Hintergrund der Merkmalsänderungen z w i s c h e n den Ausschnitten. Im übertragenen Sinne könnte man sagen, daß der "Signal-Rausch-Abstand" als Verhältnis der Merkmalsänderung aufgrund des Klicks und der Merkmalsänderungen aufgrund der Melodie innerhalb der Ausschnitte geringer als zwischen ihnen ist.

Befindet sich der Klick nun an einer Position, an der mehrere Grenzen von Binnengliedern der unterschiedlichsten Repräsentationsniveaus aufeinandertreffen, wie z.B. an der Hauptkonstituentengrenze, wo die Phrasengrenze zwischen Vorder- und Nachsatz, die Subphrasengrenze zwischen Vordersatzanschlußglied und Nachsatzmotiv, die Taktgrenze zwischen dem letzten Takt des Vordersatzes und dem ersten Takt des Nachsatzes sowie die Grenze der benachbarten Füße dieser Takte zusammenfallen, dann ist der "Signal-Rausch-Abstand" für den Klick um so günstiger, je mehr Grenzen der verschiedenen Niveaus zusammenfallen. Ein Klick an der Hauptkonstituentengrenze sollte also schneller zu entdecken sein als einer an der Grenze der beiden Subphrasen innerhalb des Vorder- oder Nachsatzes, und am längsten sollte die Detektion eines Klicks dauern, der so zwischen zwei Tönen eingebettet ist, daß keine übergeordnete Phrasengrenze mit dieser Position zusammenfällt, also z.B. ein Klick zwischen den Tönen des letzten Fußes unmittelbar am Ende des Vordersatzes. Diese Vorhersagen werden wir u.a. im Experiment 4 überprüfen.

Nachdem der Klick entdeckt worden ist, muß seine Position in Relation zu der der Töne der Melodie bestimmt werden. Um diesen Prozeß erklären zu können, müssen wir in Analogie zur Annahme von Wolff (1977), wonach die Positionen von Elementen innerhalb einer Verarbeitungseinheit auf einem Repräsentationsniveau nicht unterschieden werden können, weil sie nicht als solche, sondern als ein durch den gemeinsamen Ort definierter Merkmalsverband repräsentiert sind, annehmen, daß die Töne innerhalb der Binnenglieder auf einem Repräsentationsniveau zeitlich simultan repräsentiert sind. Die Position einzelner Töne ist dann nur durch Repräsentationsebenen-

wechsel zu ermitteln. Eine globale Analyse eines Binnengliedes auf einem bestimmten Repräsentationsniveau ist allerdings möglich, nämlich die Analyse der Binnengliederung durch Fourieranalyse zum Zwecke der Fokussierung der Binnenglieder auf der untergeordneten Repräsentationsebene. Von der übergeordneten Ebene aus betrachtet läßt sich über den Inhalt der Binnenglieder auf der untergeordneten Ebene nur soviel an Information gewinnen, daß entschieden werden kann, ob der Klick z w i s c h e n zwei Binnengliedern oder i n n e r h a l b eines Binnengliedes liegt, nicht aber die genaue Position des Klicks, wenn dieser innerhalb eines Binnengliedes liegt.

Im Prinzip könnte dieser Einschachtelungsprozeß so lange durchgeführt werden, bis der Klick an der Grenze zweier Binnenglieder lokalisiert worden ist. Die Entscheidungsstrategie des kognitiven Systems sollte darin bestehen, den Einschachtelungsprozeß dann abzubrechen, wenn bei der globalen Analyse der Binnenglieder eines untergeordneten Niveaus nicht mehr diskriminiert werden kann zwischen "der Klick liegt innerhalb eines Binnenglieds" und "der Klick liegt zwischen zwei Binnengliedern". Diese Diskrimination sollte um so schwieriger sein, je geringer der "Signal-Rausch-Abstand" für den Klick ist, d.h. je tiefer der Klick in die Merkmalsvariationen der Melodie eingebettet ist. Wenn diese Diskrimination nicht mehr möglich ist, wird der Klick an der nächsten Grenze zwischen den zuletzt fokussierten Binnengliedern eines Repräsentationsniveaus lokalisiert.

Aus diesen Überlegungen ist die folgende Hypothese abzuleiten: Die Verteilung der Klicklokalisationen sollte so aussehen, daß die Lo-

kalisationshäufigkeiten auf den einzelnen Melodiepositionen (Zwischenräumen zwischen Tönen) direkt proportional der Ordnung der Konstituentengrenzen ist, d.h. die größte Häufigkeit sollte an der Hauptkonstituentengrenze zu beobachten sein, denn dort können die Klicks exakt lokalisiert werden, aber Klicks an den unmittelbar benachbarten Positionen sollten in Richtung auf die Hauptkonstituentengrenze verschoben werden; die nächst niedrigere Häufigkeit sollte an den Positionen der nächst niedrigeren Konstituentengrenze zu beobachten sein, denn dort ist die Diskrimination des Klicks von den Merkmalsänderungen der Melodie bereits ein wenig schlechter und die Tendenz der Klicks an den unmittelbar benachbarten Positionen, in Richtung auf die Konstituentengrenze verschoben zu werden, ist gegenüber der in Richtung auf die Hauptkonstituentengrenze etwas reduziert, weil die Hauptkonstituentengrenze, die ja ebenfalls Grenze des nächst niedrigeren Niveaus ist, ebenfalls als Verschiebungsrichtung in Frage kommt, usw., schließlich sollte die Lokalisationshäufigkeit für Klicks an der niedrigsten Konstituentengrenze (Klicks zwischen den Tönen eines Fußes) die geringste sein, denn von dieser Position aus sollten beinahe alle Lokalisationen zu den nächst höheren Grenzen hin verschoben werden.

2.2.2 METHODE

Reizmaterial und Darbietungsweise

Die Melodien wurden mit Hilfe der Syntax generiert, wobei wir die Einschränkung machten, daß in allen Fällen eine Konstituentenstruktur entstehen sollte, die in einer auf jeder Strukturebene binär

verzweigten, achttaktigen Periodenstruktur besteht (vgl. Abb. 31). Diese Einschränkung war notwendig, um über alle Melodien hinweg konstante Strukturbedingungen zu haben, die einen Vergleich der Klicklokalisationen bei verschiedenen Melodien gestatten. Außerdem wurden alle Melodien mit gleich langen Notenwerten (Viertelnoten) dargeboten, damit alle Klickpositionen durch konstante Tondauern und Zeitabstände zwischen ihnen definiert waren.

Der Klick wurde immer am rechten, die Melodie am linken Ohr dargeboten. Bei dieser Anordnung erzielte Gregory (1978) die stärksten Klickverschiebungseffekte.

Die Tondauern betrugen jeweils 225 msec, das ISI ebenfalls 225 msec. Die Melodien wurden im 4/4-Takt dargeboten mit Hauptbetonung auf dem ersten und Nebenbetonung auf dem dritten Viertel. Die Lautstärke der unbetonten Töne (gemessen am Kopfhörer) betrug 7o db(A), die der Töne mit Hauptbetonung 78 db(A) und die mit Nebenbetonung 74 db(A) (Amplitudenanstieg und -abfall wie in Experiment 1).

Die Frequenzen der Töne wurden aus dem Bereich zwischen 44o und 176o Hz gewählt.

Der Klick wurde immer genau in der Mitte zwischen zwei Tönen produziert. Die Klickdauer betrug 7 msec, die Spitzenamplitude des Klicks betrug 83 db(A) (ungefähre Werte aufgrund Oszilloskopdarstellung), seine subjektive Lautstärke entsprach aber der des nebenbetonten Tons (bestimmt durch Ton-Klick-Paarvergleiche nach der Herstellungsmethode mit 3 Vpn).

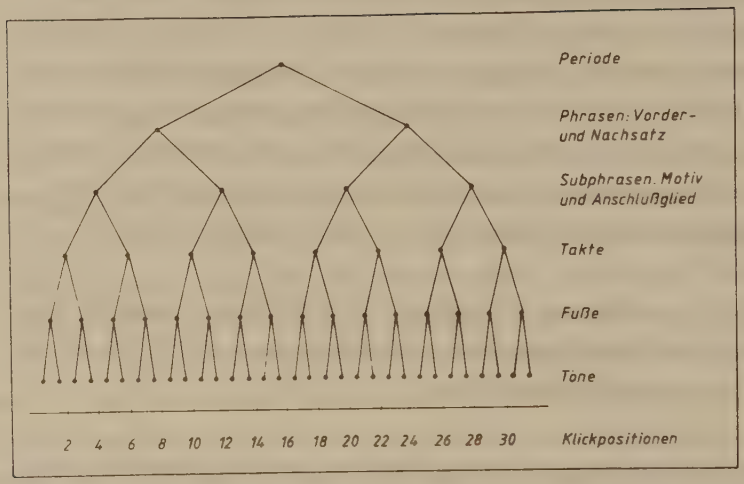


Abb. 31: Periodenstruktur der in den Experimenten 3 bis 5 dargebotenen Melodien mit der Numerierung der Klickpositionen

Versuchsplan

Von den in Abb. 31 dargestellten 31 möglichen Klickpositionen benutzten wir nur die Positionen 8 bis 24, da innerhalb dieses Bereichs alle Niveaus von Konstituentengrenzen vorkommen. Jeweils vier Melodien wurden pro Klickposition dargeboten. Hinzu kamen je zwei Melodien mit den Positionen 3 und 6 sowie 26 und 29, um die Aufmerksamkeit der Vpn nicht auf den Bereich der Positionen 8 bis 24 einzuschränken. Diese Positionen wurden in der Auswertung allerdings nicht berücksichtigt. Jede Vp mußte für jede dieser 4 (Wiederholungen) X 17 (Klickpositionen) = 68 Melodien eine Klicklokalisation vornehmen.

Versuchsapparaturen

Es wurden dieselben Geräte wie in den Experimenten 1 und 2 verwendet. Der Klick wurde durch ein schnelles Relais erzeugt, daß den rechten Kanal der Leitung zwischen Kopfhörer und Vorverstärkerausgang kurzzeitig kurzschloß. Während dieser Zeit wurde vom Tongenerator ein 3500 Hz Ton produziert, der aber aufgrund des Relaisklicks nicht mit definierter Tonhöhe wahrnehmbar war. Das Relais war mit einer Schallisolierung abgedeckt.

Versuchsdurchführung

Die Primäraufgabe der Vpn bestand in der Klassifizierung der zu hörenden Melodien in "alte Melodien" und "neue Melodien". Alte Melodien waren 17 Melodien, die in einer ersten Versuchsphase dreimal hintereinander dargeboten wurden, "neue Melodien" waren den Vpn unbekannt.

Die Sekundäraufgabe bestand in der Lokalisation der Klickposition. Dazu stand den Vpn für jede Melodie ein Blatt zur Verfügung, auf dem die Noten der jeweiligen Melodie sowie Markierungen unter allen 31 Notenzwischenräumen aufgezeichnet waren, in die die Klickposition eingetragen werden sollte. Während des Hörens durften die Vpn jedoch den Notentext nicht verfolgen. Erst nachdem sie den Klick entdeckt hatten, durften sie umblättern und die Klickmarkierung im Notentext vornehmen. Die Einhaltung dieser Vereinbarung wurde vom Vl ständig kontrolliert, der über einen Kopfhörer mithören konnte. Eine Wichtigkeitsgewichtung für Primär- und Sekundär-

aufgabe ging aus der Instruktion nicht hervor.

Die 17 alten Melodien wurden zweimal dargeboten und außerdem 34 verschiedene neue Melodien. Die Reihenfolge von alten und neuen Melodien sowie die der Klickpositionen war zufällig, aber für alle Vpn die gleiche.

Ein Durchgang mit einer Melodie wurde durch einen Warnton von 2500 Hz und von 200 msec Dauer eingeleitet. Darauf trat eine Pause von 1 sec ein, der die Melodie folgte. Der nächste Durchgang begann 20 sec nach Ende der Melodie; damit stand den Vpn ausreichend Zeit zur Verfügung, die Klickmarkierung durchzuführen und das Wiedererkennungsurteil abzugeben.

In der Instruktion wurde den Vpn erklärt, daß sie zuerst 17 Melodien dreimal hintereinander hörten, die sie sich so gut wie möglich einprägen sollten, um sie später unter 34 neuen Melodien wiedererkennen zu können. Nach einer Pause von 5 Min. begann der eigentliche Versuch. Hierzu wurde den Vpn in der Instruktion die Aufgabe gestellt, sich die Melodien aufmerksam anzuhören und sie als "alt" oder "neu" zu klassifizieren. Sobald sie auf dem rechten Ohr ein Klickgeräusch hörten, sollten sie das Blatt umblättern und im Notentext die Klickposition markieren. Erst anschliessend sollte die Klassifizierung der Melodie in "alt" oder "neu" vorgenommen werden. Sie wurden gebeten, nicht vorzeitig umzublättern.

Zur Einübung wurde den experimentellen Durchgängen 6 Übungsdurchgänge vorangestellt, in denen ausschließlich die nicht im Experi-

ment berücksichtigten Klickpositionen 1 bis 7 und 25 bis 31 benutzt wurden. Über die Einschränkung der Klickpositionen auf die Positionen 8 bis 24 wurden die Vpn nicht informiert.

Versuchspersonen

Alle 44 Vpn, die das Syntaxtraining erfolgreich absolviert hatten, nahmen auch an diesem Experiment teil. Ein Teil der Vpn wurde mit 8 DM pro Stunde bezahlt; die Psychologiestudenten erhielten stattdessen Testate für abgeleistete Vpn-Stunden.

2.2.3 BEFUNDE und DISKUSSION

Auswertungsmethode -

Die von den Vpn vorgenommenen Klicklokalisationen verrechneten wir zu zwei Werten: 1. Lokalisationshäufigkeit pro Klickposition; hierfür ist die Nullhypothese dann realisiert, wenn die Verteilung der Lokalisationshäufigkeit über die 17 Klickpositionen einer Gleichverteilung entspricht; 2. Netto-Klickverschiebungsbetrag (Klickverschiebungsscore) pro Klickposition: Berechnet wurde der Mittelwert V für die 4 Differenzen zwischen markierter Klickposition und wahrer Klickposition für jede einzelne Klickposition. Der Betrag von V gibt die Stärke der Verschiebung in Einheiten von Positionen an, sein Vorzeichen die vorherrschende Verschiebungsrichtung (positives Vorzeichen ≙ Verschiebung überwiegend in Richtung höherer Positionen; negatives Vorzeichen ≙ Verschiebung überwiegend in Richtung niedrigerer Positionen).

In die Auswertung der Klicklokalisationshäufigkeiten wurden nur solche Klickmarkierungen einbezogen, die auf den Positionen 8 bis 24 vorgenommen wurden. Von den 44 (Vpn) X 17 (Klickpositionen) X 4 (Wiederholungen) = 2992 abgegebenen Lokalisationsurteilen fielen lediglich 24 (≙ o.8 %) auf die unberücksichtigt gebliebenen Positionen. Bei der Berechnung des Klickverschiebungsscores wurden alle Lokalisationen berücksichtigt.

Sowohl für die Lokalisationshäufigkeiten als auch für die Klickverschiebungsscores rechneten wir eine einfaktorielle Varianzanalyse für abhängige Messungen. Da es sich bei den Lokalisationshäufigkeiten um abhängige Häufigkeiten handelt, ist die Überprüfung der Nullhypothese mit Chi-Quadrat-Test nicht möglich. Da die visuelle Inspektion der Verteilung der Lokalisationshäufigkeiten über alle Vpn und Klickpositionen hinweg keine auffälligen Abweichungen von einer Normalverteilung zeigte, entschlossen wir uns trotz des zu niedrigen Skalenniveaus für die Auswertung durch Varianzanalyse, verwandten aber auch hier die konservative Methode der Signifikanzniveaubestimmung nach Greenhouse & Geisser (1959). Eine Überprüfung der varianzanalytischen Ergebnisse mit dem Friedman-Test (Winer, 197o) ergab sowohl für die Variable "Lokalisationshäufigkeit" als auch für die Klickverschiebungsscores Chi-Quadrat-Werte, die auf demselben Niveau wie die varianzanalytischen F-Werte signifikant waren.

Klicklokalisationshäufigkeiten

Die Verteilung der Lokalisationshäufigkeiten über die 17 Klick-

positionen weicht sehr signifikant von einer Gleichverteilung ab ($F(43, 688) = 1o2.7$, $p<o.ool$, bei $df = 1,43$ ebenfalls $p<o.ool$; Chi-Quadrat mit $df = 16$ bei Friedman-Test: 51.6, $p<o.ool$).

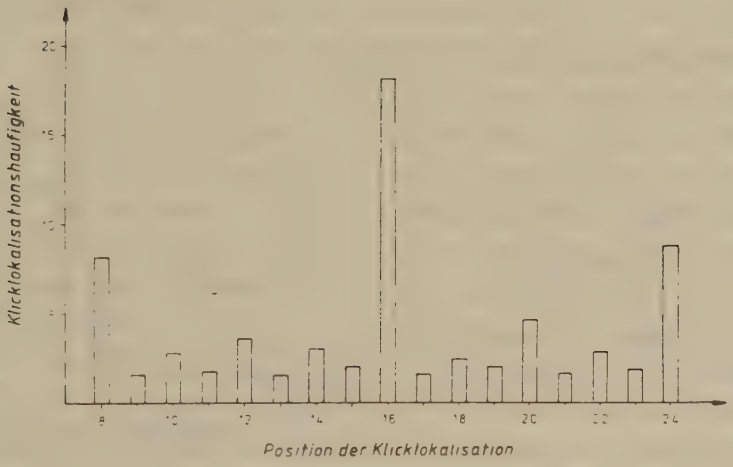


Abb. 32: Verteilung der Klicklokalisationshäufigkeiten über die Klickpositionen 8 bis 24 (gemittelt über alle 44 Vpn)

Abb. 32 zeigt eine sehr symmetrische Verteilung: Die höchste Lokalisationshäufigkeit besitzt die Position 16 (Hauptkonstituentengrenze zwischen Vorder- und Nachsatz), die zweithöchsten Häufigkeiten finden wir an den Positionen 8 (zweithöchste Konstituentengrenze der Periode zwischen Vordersatzmotiv und Vordersatzanschlußglied) und Position 24 (ebenfalls zweithöchste Konstituentengrenze

der Periode zwischen Nachsatzmotiv und Nachsatzanschlußglied). Die Mittelwertdifferenz zwischen dem Wert an Position 16 und jedem einzelnen Wert an den Positionen 8 und 24 ist sehr signifikant (Scheffé-Test, $p \leq o.ool$), die zwischen Position 8 und 24 dagegen nicht ($p \geq o.lo$).

Die dritthöchsten Lokalisationshäufigkeiten finden wir an den Positionen 12 (nächst niedrigere Konstituentengrenze zwischen den Takten des Vordersatzanschlußglieds) und an der Position 2o (zwischen den Takten des Nachsatzmotivs). Die Differenz dieser beiden Häufigkeiten ist nicht signifikant ($p \geq o.lo$), die paarweisen Differenzen zu den Werten an den Positionen 8 und 24 sind dagegen sehr signifikant ($p \leq o.ol$). Allerdings unterscheiden sich die Werte an den Positionen 12 und 2o in paarweisen Kontrasten nicht signifikant von allen übrigen, bisher noch nicht berücksichtigten Positionen, an denen sich Konstituentengrenzen noch niedrigerer Ordnung befinden ($p \geq o.o5$). Bei Scheffé-Kontrasten höherer Ordnung zeigt sich jedoch, daß die Differenz zwischen dem Mittelwert aller 8 Lokalisationshäufigkeiten innerhalb von Füßen (ungradzahlige Positionen) und dem Mittelwert der beiden Lokalisationshäufigkeiten an den Positionen zwischen Takten (Positionen 12 und 2o) auf dem o.1 %-Niveau signifikant ist.

Die Lokalisationshäufigkeiten zeigen also eine deutliche Gliederung der achttaktigen Periode im Vorder- und Nachsatz (Grenze an Position 16) und deren Untergliederung in Motive und Anschlußglieder (Positionen 8 und 24), eine Ebene einzelner Takte (Taktgrenze an den Positionen 8, 12, 16, 2o und 24), sowie deren Untergliederung in einzelne Töne. Die Lokalisationshäufigkeiten spiegeln

damit exakt die syntaktisch definierte Periodenstruktur wider, bis auf die Ebene der Füße, die zwischen denen der Takte und Töne liegt. Die Differenzierung der Fußebene deutet sich zwar in den Lokalisationshäufigkeiten an, erreicht aber nicht die Signifikanzschwelle.

Klickverschiebungsscores

Abb. 33 zeigt die Mittelwerte der Klickverschiebungsscores in Abhängigkeit von der wahren Klickposition. Ein hoher positiver Wert an einer Klickposition bedeutet, daß die an dieser Position befindlichen Klicks in Richtung einer höheren Klickposition lokalisiert wurden, ein negativer Wert besagt, daß eine Verschiebung in Richtung niedrigerer Klickposition vorgenommen wurde. Der in Abb. 33 dargestellte Kurvenverlauf weicht sehr signifikant von einem abszissenparallelen Verlauf ab ($F(43, 688) = 21.04$; $p<0.001$; bei df = 1,43 ebenfalls $p<0.001$; Chi-Quadrat mit df = 16 bei Friedman-Test: 38.4, $p<0.005$).

Die Differenz zwischen allen positiven und allen negativen Verschiebungsmittelwerten beträgt 0.12 und ist nicht signifikant von Null verschieden (Scheffé-Test, $p>0.10$). Weder die bei sprachlichem Material gefundene syntaxunabhängige Vorverlagerung des Klicks noch die von Gregory (1978) beobachtete Tendenz zu einer verspäteten Lokalisation kann durch unsere Daten bestätigt werden.

Im Unterschied zur Verteilung der Lokalisationshäufigkeiten zeigt die Kurve der Klickverschiebungsscores, durch welche Verschiebungs-

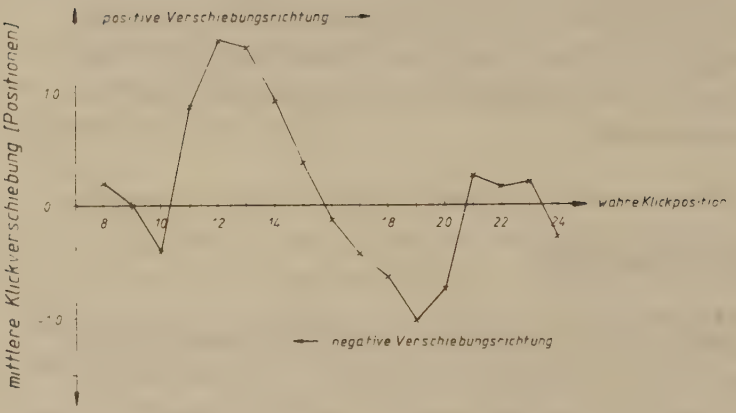


Abb. 33: Mittlere Klickverschiebungsscores in Abhängigkeit von der wahren Klickposition

tendenzen die beobachteten Unterschiede zwischen den Lokalisationshäufigkeiten zustande gekommen sind. Diese Information ist insofern von Bedeutung, als die Hypothese über die Fehllokalisationen ja auch Aussagen über die Richtung der Verschiebungstendenzen macht. Die Stärke der Verschiebungstendenzen sollte von der Tiefe der Einbettung des Klicks in die Variationsstruktur der Melodie abhängen (je niedriger die Konstituentengrenze, an der sich der Klick befindet, desto höher die Tendenz zur Verschiebung), aber sie wird durch den Abstand der Klickpositionen zur nächst höheren Konstituentengrenze begrenzt.

Aus dieser Hypothese ist abzuleiten, daß eine Verschiebungsten-

denz am stärksten in der Nähe der Hauptkonstituentengrenze auftreten und an den zweithöchsten Konstituentengrenzen schwächer als in der Nähe der Hauptkonstituentengrenze sein sollte. Abb. 33 zeigt, daß diese Vorhersage mit den Daten übereinstimmt: Zwischen Positionen 12 und 16 nimmt der mittlere Verschiebungsbetrag ab, nimmt aber im negativen Bereich zwischen den Positionen 16 und 19 wieder zu (die Differenz zwischen den beiden Extremwerten an Position 12 und 19 ist sehr signifikant, Scheffé-Test, $p<0.001$). Außerdem ist der Verschiebungsbetrag in Richtung auf die Hauptkonstituentengrenze an Position 16 größer als in Richtung der nächst niedrigeren Konstituentengrenzen an den Positionen 8 und 24 (die Differenz zwischen dem Mittelwert aller Verschiebungsscoremittelwerte an den Positionen 8 bis 10 und dem Verschiebungsscoremittelwert an Position 12 ist signifikant, Scheffé-Test, $p<0.01$; ebenso die Differenz zwischen dem Mittelwert aller Verschiebungsscoremittelwerte an den Positionen 21 bis 24 und dem Verschiebungsscoremittelwert an Position 19, $p<0.05$).

Die Positionen 12 und 20 stellen die Taktgrenzen unmittelbar vor der Hauptkonstituentengrenze dar. Die positive Klickverschiebung in Richtung der Hauptkonstituentengrenze beginnt bereits eine Position vorher, d.h. die Hauptkonstituentengrenze übt ihre "Anziehungskraft" bereits 1 1/4 Takte vorher aus. Die negative Klickverschiebung in Richtung der Hauptkonstituentengrenze reicht bis genau einen Takt hinter diese Grenze. Demgegenüber ist die Verschiebungstendenz in Richtung auf die nächst niedrigeren Konstituentengrenzen (Grenzen zwischen Motiven und Anschlußgliedern) an Position 8 und 24 bedeutend geringer ausgeprägt. Dies ist aber

nicht darauf zurückzuführen, daß an den Positionen 8 bis 1o bzw. 21 bis 24 kaum Verschiebungen stattfinden, sondern darauf, daß sich positive und negative Verschiebungsbeträge gegenseitig aufheben, weil z.B. bei einem Klick an Position 1o sowohl nach Position 8 als auch nach Position 12 verschoben werden kann, wobei die Verschiebungstendenz in Richtung auf Position 8 überwiegt, da dort die höhere Konstituentengrenze liegt. Dasselbe gilt für Klicks an den Positionen 21 und 22.

Die Repräsentationsstruktur für die dargebotenen Melodien ist bei unseren syntaktisch kompetenten Vpn also tatsächlich eine hierarchische, und diese hierarchische Struktur entspricht genau derjenigen Struktur, die die syntaktische Beschreibung der Melodien liefert. Allerdings zeigte sich auf dem Niveau der Füße im Vergleich zu dem der einzelnen Töne kein differentieller Fehllokalisationseffekt: Die Lokalisationshäufigkeiten innerhalb von Füßen waren gleich denen zwischen Füßen, d.h. unterhalb des Taktniveaus kann offenbar nicht mehr so gut wie oberhalb des Taktniveaus zwischen den Klickpositionen differenziert werden.

Dieses Ergebnis ist im Rahmen unserer Erklärung des Fehllokalisationseffekts dahingehend zu interpretieren, daß für Klicks, die sehr tief in die Variationsstruktur der Merkmale einer Melodie eingebettet sind, eine Diskrimination zwischen "Klick liegt innerhalb eines Fußes" und "Klick liegt zwischen zwei Füßen" bei Erreichen der Fußebene im Verlauf des hierarchischen Einschachtelungsprozesses zum Auffinden der Klickposition nicht mehr möglich ist. Die Merkmalsvariation innerhalb eines Taktes ist zu groß, als daß der Variationsanteil, der durch den Klick erzeugt wird, sich hiervon

abheben könnte. Die gleich schlechte Lokalisation von Klicks an den beiden untersten Konstituentengrenzen ist deshalb nicht notwendigerweise darauf zurückzuführen, daß die Repräsentanzen der Takte keine Binnengliederung in Füße aufweisen würde, sondern nur eine in einzelne Töne. Wir neigen hier eher zu der Interpretation, daß die Verarbeitung der Informationen, die Repräsentanzen unterhalb des Taktniveaus zu entnehmen sind, aufgrund einer relativ zu übergeordneten Repräsentationsniveaus zu großen Merkmalsvariation weniger detailliert und eher global ablaufen muß. Dies gilt aber nur für Taktrepräsentanzen, die in eine längere und mehrstufig hierarchisch organisierte Melodie eingebettet sind, und nicht für entsprechend kurze Tonfolgen ohne Kontext, wie die durchaus detaillierten Identifikationsleistungen in den Experimenten 1 und 2 belegen. Der phänomenale Eindruck der Ganzheitlichkeit der Melodiewahrnehmung rührt demnach teilweise daher, daß Informationen über die Binnengliederung der Takte nur sehr global sein können, wenn eine differenzierte und variationsreiche musikalische Struktur wahrgenommen wird.

2.3 EXPERIMENT 4: VERGLEICH SYNTAKTISCH TRAINIERTER UND UNTRAINIERTER HÖRER IN EINER KLICKDETEKTIONSAUFGABE

2.3.1 EINLEITUNG

In der Einleitung zu Experiment 3 haben wir bereits die Hypothese aufgestellt und begründet, daß ein Klick in einer hohen Konstituentengrenze schneller als einer in einer niedrigeren Konstituentengrenze zu entdecken sein sollte. Diese Hypothese wurde mit

sprachlichem Material durch die Reaktionszeitdaten von Abrams & Bever (1969) nicht gestützt, während Holmes & Foster (197o) mit der Hypothese eindeutig konforme Befunde erhielten. Diese Hypothese soll im folgenden Experiment mit musikalischem Material überprüft werden.

Den Befund aus Experiment 3, daß die Klicklokalisationshäufigkeiten auf strukturellen Niveaus unterhalb von Takten nicht zwischen dem Fußniveau und dem Tonniveau differenzieren, haben wir auf Diskriminationsprobleme zurückgeführt, die auf strukturell niedrigen Repräsentationsniveaus während des hierarchischen Einschachtelungsprozesses auftreten, wenn darüber zu entscheiden ist, ob auf dem gerade zu analysierenden Repräsentationsniveau der Klick innerhalb eines Binnengliedes dieses Niveaus oder zwischen zwei Binnengliedern liegt. Wenn vom Hörer allerdings keine exakte Lokalisation, sondern nur eine Reaktion auf die Detektion des Klicks verlangt wird, sollten diese Diskriminationsprobleme nicht auftreten, d.h. die Detektionszeit für Klicks zwischen den Tönen eines Fußes sollte höher als die für Klicks zwischen Füßen sein.

Wir haben bisher den Funktionen kognitiver Strukturen bei der Wahrnehmung wenig Beachtung geschenkt und lediglich den Beitrag kognitiver Schemata oder Frames im Zusammenhang mit strukturellen Identifikationsprozessen beschrieben, wie wir sie in den Experimenten 1 und 2 untersucht haben. Diese Vernachlässigung ist eine - allerdings nicht notwendige - Konsequenz der Gibson'schen Wahrnehmungstheorie (Gibson, 1966a), die einen Schwerpunkt bei der Abhängigkeit der Wahrnehmung von den Strukturen der Objekte und der

Umwelt setzt und dabei den Anteil des Wahrnehmenden so gut wie vollständig ignoriert:

"Nevertheless, the Gibsonian view of perception also seems inadequate (gemeint ist mit "also" die Sichtweise der konstruktivistischen Informationsverarbeitungsmodelle, Anm. des Verf.), if only because it says so litte about the perceiver's contribution to the perceptual act. There must be definite kinds of structure in every perceiving organism to enable it to notice certain aspects of the enviroment rather than others, or indeed to notice anything at all" (Neisser, 1976, S. 9).

Solche Wahrnehmungssituationen, wie das Suchen nach einem bestimmten Motiv in einer Sinfonie, von dem man erwartet, es müsse jeden Augenblick kommen, oder das selektive Beachten des Rhythmus oder anderer musikalischer Strukturaspekte, setzen beim Hörer eine bestimmte Intention voraus, deren Realisierung während eines Wahrnehmungsvorgangs nur dann möglich ist, wenn das gesuchte Motiv bereits in abstrakter Form mehr oder weniger detailliert repräsentiert ist oder der Hörer beim selektiven Beachten des Rhythmus eine abstrakte Vorstellung darüber besitzt, was einen Rhythmus als Wahrnehmungsinhalt auszeichnet.

Bevor das gesuchte Motiv auftritt und während der Hörer den Rhythmus beachtet, finden nicht notwendigerweise strukturelle Identifikationsprozesse statt, und doch ist der Hörer keineswegs inaktiv, sondern wählt unter den sich bildenden Repräsentationen vom Wahrnehmungsobjekt aus. Neisser (1976) sagt, das Wahrnehmungsobjekt werde exploriert, und diese Exploration, die aktive Suche und Aufnahme der den Intentionen des Wahrnehmenden entsprechenden Repräsentationen des Wahrnehmungsobjekts, werde durch antizipa-

torische Schemata geleistet. Den ständig erneut zu durchlaufenden Zyklus von Antizipation, Exploration und Modifikation der Antizipation aufgrund der wahrgenommenen Informationen über das Wahrnehmungsobjekt bezeichnet er als "perceptual circle".

Die beiden Prozesse, die wir bisher im Rahmen des Hierarchiemodells beschrieben haben, nämlich die Fourieranalyse und die Invarianzanalyse bei der strukturellen Identifikation, sind zwar Teilprozesse des Wahrnehmungszyklus, beschreiben ihn aber keineswegs funktional vollständig. Die Fourieranalyse ist lediglich der Mechanismus, der den selektiven Zugriff auf die Binnenglieder einer Repräsentationsebene vermittelt. Es handelt sich hierbei nicht um einen kontrollierten, sondern um einen automatisch ablaufenden Prozeß, der lediglich initialisiert zu werden braucht. Die Initialisierung aber muß in der Kontrolle des Wahrnehmenden liegen, der über die Richtung des intendierten Repräsentationsebenenwechsels zu entscheiden und das Ergebnis der Fourieranalyse, die Binnengliederung einer Repräsentationsebene, auf seine Brauchbarkeit in bezug auf seine Intention zu überprüfen hat. Solche Aktivitäten meint Neisser, wenn er von antizipatorischen Schemata spricht, die die Aufnahme von Informationen steuern.

Andererseits ist der beschriebene Identifikationsprozeß einer, der erst dann stattfinden kann, wenn die antizipierte, den Intentionen des Wahrnehmenden entsprechende Information aufgenommen worden ist. Die Entscheidung darüber, ob diese Information den Spezifikationen genügt, die über sie im Wahrnehmungsschema enthalten sind, und ob eine Identifikation eingeleitet werden soll, ist ebenfalls eine Funktion der Wahrnehmungsschemata, die bereits einiges über die

Objekte der Wahrnehmungswelt "wissen" müssen, um diese Steuerungsfunktionen übernehmen zu können.

Inhalt solcher Wahrnehmungsschemata könnte z.B. die Spezifikation sein, welches Repräsentationsniveau aufgrund der Intention des Hörers angesteuert werden muß, um eine bestimmte Identifikation leisten zu können. Oder das Wahrnehmungsschema könnte eine Erwartung über das bevorstehende Ende eines Musikstücks oder einer Passage daraus enthalten, weil es die noch zu erwartende ungefähre Anzahl von Binnengliedern des gerade fokussierten Niveaus enthält.

Für die visuelle Wahrnehmung und die alltägliche auditive Wahrnehmung von Schallquellen besitzen diese Schemata sicherlich Orientierungsfunktionen: Der Wahrnehmende weiß jederzeit, wo in seiner Umwelt und in welcher Art von Umwelt er sich gerade befindet, ohne hierzu ständig eine detaillierte Identifikation aller Umweltobjekte vornehmen zu müssen. Die gerade genannten Beispiele für die durch Wahrnehmungsschemata geleisteten Spezifikationen über die aufzunehmenden Informationen sind ebenfalls Beispiele für eine Art "musikalischer Orientierung". Der musikalisch erfahrene Hörer weiß zumindest sehr global aufgrund seines impliziten Wissens über Strukturen in Musik, wo er sich gerade innerhalb des musikalischen Strukturgebäudes befindet, dies auch dann, wenn er die Musik erstmalig hört, vorausgesetzt allerdings, er verfügt über ausreichende Erfahrungen mit dem musikalischen Stil und dem musikalischen Formtyp.

Der musikalisch erfahrene Hörer strukturiert die zu hörende Musik kognitiv, indem er ständig Erwartungen und Wahrnehmungen miteinander in Beziehung setzt (vgl. Meyer, 1956), der musikalisch wenig erfahrene Hörer überläßt sich eher passiv der Wahrnehmung der Variationsstrukturen. Auf die Melodien unserer Syntax und auf unsere trainierten Hörer bezogen erwarten wir von ihnen, daß sie versuchen genau zu wissen, wo innerhalb der Periodenstruktur sie sich gerade befinden, während dies untrainierte Hörer nicht zu tun vermögen.

Stellen wir nun eine Klickdetektionsaufgabe, dann sollten die trainierten Hörer in der Lage sein, in höherem Maße ihre Aufmerksamkeit auf mögliche Klickpositionen zu lenken, weil sie sie antizipieren können, als untrainierte Hörer. Dies sollte dazu führen, daß trainierte Hörer an den besonders ausgezeichneten Stellen der Periodenstruktur, z.B. bei einem Klick an der Hauptkonstituentengrenze, schneller zu reagieren vermögen als untrainierte Hörer. Z.B. sollte für beide Gruppen die kürzeste Detektionszeit die für einen Klick an der Hauptkonstituentengrenze sein, aber im Vergleich zu den übrigen Klickpositionen sollten die trainierten Hörer hier schneller als die untrainierten reagieren können.

2.3.2 METHODE

Reizmaterial, Versuchsapparaturen und Darbietungsweise

Im Unterschied zu Experiment 3 haben wir hier nur viertaktige Melodien verwendet, die syntaktisch aus einem Vordersatz bestehen,

der allerdings immer auf dem Grundton endet. Die hierdurch ermöglichte Zeitersparnis bei der Melodiedarbietung war notwendig, um pro Klickposition möglichst viele Meßwiederholungen vornehmen zu können, ohne gleichzeitig die Versuchszeit über eine Stunde ausdehnen zu müssen. Ansonsten wurden die Melodien genauso konstruiert, wie in Experiment 3 beschrieben.

Die Darbietungsweise war identisch mit der in Experiment 3. Im Unterschied zu Experiment 3 wurde allerdings keine Klickmarkierung von den Vpn verlangt, sondern eine Reaktion auf das Entdecken des Klicks durch Druck auf eine Morsetaste.

Versuchsplan und Versuchspersonen

Klicks wurden an allen 16 möglichen Positionen zwischen Tönen dargeboten (Positionen 1 bis 16 in Abb. 31). Pro Klickposition wurden zehn Melodien verwendet, insgesamt also 160 verschiedene Melodien. Die experimentelle Gruppe bestand aus den Vpn, die das Syntaxtraining absolviert hatten, die Kontrollgruppe wurde aus Vpn zusammengestellt, die weder am Syntaxtraining teilgenommen hatten, noch ein Musikinstrument spielten. Die Stichprobe der experimentellen Gruppe umfaßte 40 Vpn, die der Kontrollgruppe 47. Die Vpn wurden mit 8 DM bezahlt mit Ausnahme der Psychologiestudenten, die Testate für die abgeleisteten Vpn-Stunden erhielten.

Dieses Experiment besteht also in einem zweifaktoriellen Versuchsplan mit unabhängigen Stichproben (experimentelle vs. Kontrollgruppe) und 16 abhängigen Bedingungen (Klickpositionen).

Versuchsdurchführung

Alle Klickpositionen wurden in zufälliger Reihenfolge dargeboten und vor jeder Versuchssitzung neu gemischt. Zu Beginn des Versuchs wurden 32 Melodien als Übungsbeispiele dargeboten, in denen alle Klickpositionen zweimal vorkamen. Den Melodien ging ein 2oo msec dauernder Warnton von 25oo Hz voraus, dem eine Pause von einer sec. folgte. Der Zeitabstand zwischen Ende einer Melodie und dem Warnton vor der folgenden Melodie betrug 6 sec.

Die Instruktion forderte die Vpn auf, die Melodien aufmerksam anzuhören, da sie einen Teil von ihnen anschließend unter noch nicht gehörten Melodien wiedererkennen sollten. Dieser Wiedererkennungstest wurde aber nicht durchgeführt. Als Sekundäraufgabe wurde die Detektion des Klicks verlangt. Die Vpn sollten so schnell wie möglich die Morsetaste drücken, sobald sie den Klick am rechten Ohr wahrgenommen hatten. In der Instruktion wurde keine Wichtigkeitsgewichtung für die beiden Aufgaben vorgenommen. Der Versuch dauerte durchschnittlich eine Stunde einschließlich einer Pause von lo Min.

2.3.3 BEFUNDE UND DISKUSSION

Auswertungsmethode

Die lo Meßwiederholungen pro Klickposition wurden gemittelt und varianzanalytisch ausgewertet. Wir rechneten eine 2 (experimentelle vs. Kontrollgruppe) X 16 (Klickpositionen) - Varianzanalyse nach der "least-squares"-Methode (vgl. Winer, 197o).

Haupteffekte

Die Differenz zwischen den RT-Mittelwerten der experimentellen Gruppe (27o msec) und der Kontrollgruppe (289 msec) ist signifikant ($F(1,85) = 4.73$, $p \leq o.o5$). Da Experiment 4 auch für die Vpn der experimentellen Gruppe der erste Versuch war, in dem sie eine Reaktion durch Tastendruck abgeben mußten, ist die höhere Reaktionsgeschwindigkeit nicht auf Übungsunterschiede in der Reaktionsabgabe zurückzuführen. Da die experimentelle Gruppe aber bereits vorher am Experiment 3 teilgenommen hatte, kann diese RT-Differenz auf die größere Erfahrung der Vpn der experimentellen Gruppe im Entdecken von Klicks auf dem Hintergrund der Merkmalsvariationen in den Melodien zurückgeführt werden. Der Unterschied geht also auf eine höhere Verarbeitungsgeschwindigkeit bei den Vpn der experimentellen Gruppe im Vergleich zu denen der Kontrollgruppe zurück und nicht auf eine größere Geschwindigkeit der motorischen Ausführung der Reaktion.

Die Reaktionszeitmittelwerte für den Haupteffekt "Klickposition" gibt Abb. 34 wieder (mittlere stark eingezeichnete Kurve); dieser Haupteffekt ist sehr signifikant ($F(15, 1275) = 89.92$, $p \leq o.ool$, bei $df = 1,86$ ebenfalls $p \leq o.ool$).

Der in Abb. 34 dargestellte Kurvenverlauf für den Haupteffekt "Klickposition" zeigt zwei Charakteristika: 1. einen Abfall der mittleren RT von der ersten zur letzten Klickposition und 2. einen nicht monotonen Verlauf, der sich als einigermaßen regelmäßiges Zickzackmuster zeigt, das besonders deutlich an den ersten 9 Positionen in Erscheinung tritt.

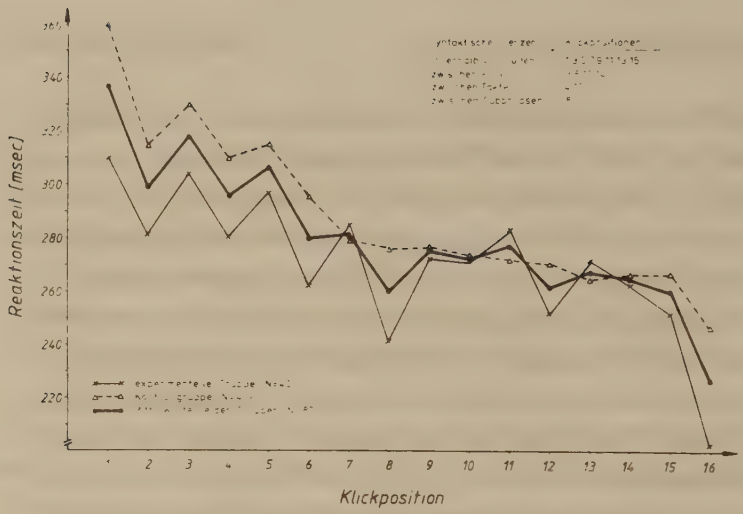


Abb. 34: RT-Mittelwerte für das Entdecken eines Klicks in viertaktigen Melodien in Abhängigkeit von der Klickposition

Einen ähnlichen Abfall der RT von der ersten zur letzten Klickposition fanden auch Holmes & Foster (197o). Sie führten ihn auf eine Abnahme der Kapazitätsbelastung bei der Verarbeitung von Sätzen zurück, da durch die jeweils vorausgegangene syntaktische Struktur die nachfolgenden Satzteile um so zutreffender antizipiert werden können, je später der Klick auftritt, d.h. je vollständiger der Satz bereits verarbeitet worden ist. Hierdurch sollte es den Vpn möglich sein, ihre Aufmerksamkeit um so mehr dem Kanal zuzuwenden, an dem der Klick auftritt, je vollständiger der Satz bereits dargeboten wurde, wenn der Klick auftritt. Eine solche Änderung der Aufmerksamkeitsverteilung könnte auch in unserem Experiment zu diesem beobachteten Effekt geführt haben, denn

die Vpn könnten die Strategie verfolgt haben, nur den Anfang einer Melodie aufmerksam zu verfolgen, in der Hoffnung, im Wiedererkennungstest die Melodien bereits aufgrund der ersten Takte wiedererkennen zu können. Für unsere Fragestellungen sind allerdings die RT-Differenzen zwischen Klickpositionen von größerer Bedeutung, die mit Konstituentengrenzen unterschiedlicher Höhe zusammenfallen. Wir rechneten Scheffé-Kontraste höherer Ordnung für die RT-Differenzen an Positionen innerhalb von Füßen (alle 8 ungradzahligen Positionen), zwischen Füßen (Positionen 2, 6, 1o und 14), zwischen Takten (Positionen 4 und 12) und zwischen Subphrasen (Position 8). Tab. 1 faßt die Ergebnisse zusammen.

Tab. 1: Irrtumswahrscheinlichkeiten und RT-Mittelwerte für die Scheffé-Kontraste höherer Ordnung für Klickpositionen an syntaktischen Grenzen unterschiedlicher Höhe

syntaktische Grenze	RT-Mittelwerte (in msec.)	Irrtumswahrscheinlichkeiten (p)		
		ZF	ZT	ZS
innerhalb von Füßen (IF)	289	$<$o.o5	$<$o.o5	$<$o.o1
zwischen Füßen (ZF)	278		$>$o.1o	$<$o.o5
zwischen Takten (ZT)	279			$<$o.o5
zwischen Subphrasen (ZS)	259			

Entsprechend unserer Hypothese zeigt sich hier also die Tendenz, daß die RTs mit zunehmender Höhe der syntaktischen Grenze, an der

sich ein Klick befindet, kürzer werden. Eine Ausnahme von dieser Tendenz bildet überraschenderweise der Vergleich der RT-Mittelwerte für Klicks zwischen Füßen und Klicks zwischen Takten, der als einziger Scheffé-Kontrast nicht die 5 %-Signifikanzschwelle unterschreitet. Dagegen differenzieren die RT-Mittelwerte für Klicks zwischen Füßen und innerhalb von Füßen, während die Lokalisationshäufigkeiten in Experiment 3 hier keine Differenzierung offenbarten. Wenn man von dieser Ausnahme einmal absieht, dann wird unsere Hypothese in Übereinstimmung mit der Tendenz, die sich in dem Klickdetektionsexperiment von Holmes & Foster (197o) zeigte, gestützt. Da die Detektion der Klicks im Unterschied zur Lokalisationsaufgabe unabhängig von einer Speicherung des tonalen Kontextes, in dem der Klick auftrat, im Kurzzeitgedächtnis möglich ist und wir auch unter diesen experimentellen Bedingungen von der syntaktischen Struktur der Melodien in systematischer Weise abhängige Befunde erhalten haben, muß die Wahrnehmung zu einer Repräsentation der Melodiestrukturen geführt haben, die in gleicher Weise hierarchisch gegliedert ist, wie die sich bei der syntaktischen Analyse der Melodien herausstellende Struktur.

Interaktion zwischen Vpn-Gruppen und Klickpositionen

Bei Zugrundelegung des konservativen Signifikanzkriteriums nach Greenhouse & Geisser (1959) ist die Interaktion zwischen Vpn-Gruppen und Klickpositionen nicht signifikant ($F(15, 1275) = 1.95$, $p < o.o2$, bei $df = 1,86$ jedoch $o.25 < p < o.1o$). Aus diesem Grunde betrachten wir die im folgenden dargestellten Unterschiede zwi-

schen den Vpn-Gruppen in Abhängigkeit von der Klickposition lediglich als Tendenzen und wählen als Kriterium für eine Tendenz bei Scheffé-Kontrasten das 10 %-Niveau.

Wir berechneten für alle konsekutiven Paare von Klickpositionen Scheffé-Kontraste zwischen der Differenz der RT-Mittelwerte bei der Kontrollgruppe an Position i und i+1 ($k_i - k_{i+1}$) und der Differenz der RT-Mittelwerte bei der experimentellen Gruppe an diesen Positionen ($e_i - e_{i+1}$), so daß sich der Kontrast

$$\frac{1}{2}(k_i - k_{i+1}) - \frac{1}{2}(e_i - e_{i+1}) = \frac{1}{2}(k_i + e_{i+1}) - \frac{1}{2}(e_i + k_{i+1})$$

ergibt, der die Parallelität des Kurvenverlaufs beider Stichprobenmittelwerte an den Positionen i und i+1 überprüft.

Die RT-Mittelwerte für beide Gruppen getrennt sind ebenfalls in Abb. 34 dargestellt. Nach dem vereinbarten Tendenzkriterium sind die folgenden partiellen Interaktionen bedeutsam:

1. Von Position 6 (Klick zwischen Füßen des 2. Taktes) nach 7 (Klick innerhalb des 2. Fußes im 2. Takt) ist bei der experimentellen Gruppe ein hypothesenkonformer RT-Anstieg und bei der Kontrollgruppe eine RT-Abnahme zu beobachten ($p \leq 0.05$).

2. Von Position 7 (Klick innerhalb des 2. Fußes im 2. Takt) nach 8 (Klick an Subphrasengrenze) ist bei der experimentellen Gruppe eine stärkere hypothesenkonforme RT-Abnahme als bei der Kontrollgruppe zu beobachten ($p \leq 0.05$).

3. Von der Position 8 (Klick an Subphrasengrenze) nach 9 (Klick innerhalb des 1. Fußes des 3. Taktes) steigt bei der experimentellen Gruppe die RT der Hypothese entsprechend an, während sie

bei der Kontrollgruppe konstant bleibt ($p \leq o.lo$).

4. Von Position 11 (Klick innerhalb des 2. Fußes des 3. Taktes) nach 12 (Klick an Taktgrenze zwischen 3. und 4. Takt) fällt die RT bei der experimentellen Gruppe wie erwartet ab, aber stärker als bei der Kontrollgruppe, bei der sie beinahe konstant bleibt ($p \leq o.lo$).

Bis zur Position 6 einschließlich unterscheidet sich das Detektionsverhalten beider Vpn-Gruppen überhaupt nicht. Im Bereich der höchsten Konstituentengrenze (Subphrasengrenze an Position 8) und im Bereich der zweithöchsten Konstituentengrenze dahinter (Taktgrenze an Position 12) treten der Tendenz nach Unterschiede auf, die darin bestehen, daß die Kontrollgruppe im Gegensatz zur experimentellen Gruppe keinerlei Differenzierung in der Detektionsgeschwindigkeit zwischen Klicks an unterschiedlich hohen Konstituentengrenzen zeigt. Erst in der zweiten Hälfte der Melodien offenbaren sich die von uns vorhergesagten Unterschiede zwischen den beiden Gruppen: Die Vpn der experimentellen Gruppe reagieren bei Klicks an höheren Konstituentengrenzen schneller als die der Kontrollgruppe, aber zum Teil auch bei Klicks an den unteren Konstituentengrenzen langsamer als die der Kontrollgruppe (vgl. die Positionen 11 und 13 in Abb. 34).

Wir sehen in der sich abzeichnenden Tendenz die Auswirkungen einer aktiven Antizipation der jeweils nächsten Klickposition, die von beiden Vpn-Gruppen zu Beginn einer Melodie in gleicher Weise vorgenommen werden kann. Aber je mehr Strukturniveaus im Verlauf der Entwicklung der Melodie hinzukommen, desto schwieriger

wird es für die untrainierten Vpn, sich in der Struktur der Melodien zu orientieren, sie verlieren sozusagen die Übersicht, weil ihre Wahrnehmungsschemata nicht differenziert genug sind, um bei zunehmender struktureller Komplexität noch eine aktive Strukturierung und Erwartungsbildung vorzunehmen.

2.4 EXPERIMENT 5: SAME-DIFFERENT-RT BEI VERGLEICH SYNTAKTISCH DEFINIERTER MELODIEAUSSCHNITTE IN ABHÄNGIGKEIT VON IHRER SYNTAKTISCHEN UND ZEITLICHEN NÄHE

2.4.1 EINLEITUNG

Die Überzeugungskraft der Befunde zur Klicklokalisation und -detektion steht und fällt mit der Adäquatheit der Erklärung des Lokalisations- und Detektionsprozesses, die wir unter Zuhilfenahme von Zusatzannahmen aus dem Hierarchiemodell abgeleitet haben. Diese Zusatzannahmen, z.B. die über das Diskriminationsproblem, auf niedrigen Repräsentationsniveaus zwischen klickbedingter und melodiebedingter Merkmalsvariation zu unterscheiden, stellen keine zwingenden Ableitungen aus dem Hierarchiemodell dar. Die Interpretation der Befunde aus den beiden Klickexperimenten als Anzeichen für eine hierarchische Repräsentationsstruktur beim Hören von Melodien mit einer syntaktisch als Hierarchie beschriebenen Periodenstruktur ist nur dann logisch gerechtfertigt, wenn wir aus guten Gründen davon ausgehen können, daß die diese Interpretation erst ermöglichenden Zusatzannahmen psychologisch plausibel sind. Deshalb wäre es wünschenswert, Hypothesen aus dem Hierarchiemodell abzuleiten und zu überprüfen, die keine neuen Zusatzannahmen er-

fordern und auf diese Weise eine direktere Überprüfung der Hypothese einer hierarchischen Repräsentationsstruktur gestatten.

Eine solche direktere Überprüfung ist dann möglich, wenn wir Hypothesen über solche Verarbeitungsprozesse auf der Grundlage des Hierarchiemodells ableiten, die auf unterschiedliche Repräsentationsniveaus zugreifen sollten, und diese in Form von Vorhersagen über die Verarbeitungsgeschwindigkeit in Abhängigkeit von der Höhe des Repräsentationsniveaus formulieren, auf dessen Repräsentanzen zugegriffen werden sollte. Gleichzeitig können wir hierdurch weitere Informationen zur Beurteilung der kognitiven Adäquatheit gewinnen, wenn wir wieder von der Hypothese ausgehen, daß die Repräsentationen der Melodien dieselbe Struktur besitzen, wie sie sich in der syntaktischen Beschreibung zeigen.

Ein Argument für die Annahme einer hierarchischen Form der Repräsentation bestünde in dem Nachweis, daß die aus der Hypothese einer zeitlich s e r i e l l e n Organisation ableitbaren Vorhersagen über den Zugriff auf die Repräsentationsstruktur nicht zutreffen, während die aus der Hypothese einer h i e r a r c h i s c h e n Organisation ableitbaren Vorhersagen mit den Daten gut zu vereinbaren sind.

Ein Modell der auditiven Informationsverarbeitung, das von einer zeitlich seriellen Organisation der dem echoischen Speicher zu entnehmenden Repräsentanzen ausgeht, ist das Modell von Massaro (1975, 1976a). Allerdings nimmt Massaro (1977) nun nicht mehr an, der echoische Speicher enthalte jeweils nur die Information über einen einzigen Ton, wie noch Idson & Massaro (1976), sondern hat

inzwischen auch die Möglichkeit der Gruppierung von Tönen in sein Verarbeitungsmodell integriert. Aber diese Gruppen von Tönen werden letztlich doch wieder als zeitlich seriell organisiert betrachtet, so daß nach Massaro (1977) allenfalls zwei Repräsentationsniveaus vorliegen können, nämlich das der einzelnen Töne und das ihnen übergeordnete Niveau der Tongruppen.

Bei zeitlich serieller Organisation sollte das Erkennen einer strukturellen Beziehung zwischen zwei Tongruppen (z.B. in einer same/different-Aufgabe) nur vom zeitlichen Abstand und eventuell noch von der gruppeninternen Struktur und der Form der Beziehung zwischen den Gruppen abhängen. Die Verarbeitungsabfolge könnte nach Massaro's Modell dann folgendermaßen beschrieben werden: Die erste aufgabenrelevante Gruppe wird als solche identifiziert und in den "synthesized auditory store" (Massaro, 1975; modalitätsspezifischer Kurzzeitspeicher zwischen "preperceptual auditory store" und "generated abstract memory", letzterer ist dem "short term memory" vergleichbar) übernommen, wo sie einem zeitabhängigen Decay und/oder einer Interferenzwirkung aufgrund der folgenden Töne oder Tongruppen unterliegt. Wenn ebenfalls die zweite Vergleichsgruppe dort gespeichert wurde, kann der Vergleich durchgeführt werden, was um so schneller möglich sein sollte, je weniger die erste Gruppe während der Speicherung beeinträchtigt wurde.

Bei Zugriff auf eine hierarchische Repräsentationsstruktur sollte die für einen Vergleich zweier Melodieausschnitte erforderliche Verarbeitungszeit nicht nur von der zeitlichen Nähe der Vergleichsausschnitte sondern auch von der Einbettungsstruktur dieser Aus-

schnitte in übergeordnete Repräsentanzen abhängen. Gehören die Vergleichsausschnitte auf dem unmittelbar übergeordneten Repräsentationsniveau zur selben Repräsentanz, dann kann der Vergleich direkt durch Entnahme der auf dem ihnen übergeordneten Niveau repräsentierten Beziehungsstruktur vorgenommen werden. Gehören die Vergleichseinheiten aber trotz gleicher zeitlicher Nähe, wie im ersten Fall, auf dem unmittelbar übergeordneten Niveau zu verschiedenen Repräsentanzen, dann muß erst das Repräsentationsniveau fokussiert werden, auf dem die den Ausschnitten übergeordneten Einheiten selbst wieder auf dem ihnen übergeordneten Niveau zur selben Repräsentanz gehören. In diesem Fall wäre also die Beziehung der Vergleichsausschnitte nur über ein höheres Repräsentationsniveau zu entnehmen als im ersten Fall.

Ein Beispiel für den ersten Fall ist der Vergleich der beiden ersten Takte des Vordersatzmotivs, ein Beispiel für den zweiten Fall bei gleicher zeitlicher Entfernung der Vergleichseinheiten ist ein Vergleich des zweiten Taktes des Vordersatzmotivs mit dem ersten Takt des Vordersatzanschlußglieds, wobei die Beziehung dieser Takte sich erst über die Beziehung zwischen Motiv und Anschlußglied des Vordersatzes erschließt.

Auch dann, wenn zwei zeitlich nicht benachbarte Ausschnitte zu vergleichen sind, z.B. bei Vergleich des ersten Taktes des Vordersatzmotivs mit dem ersten Takt des Vordersatzanschlußglieds, ist deren Beziehung nur über diejenige Repräsentationsebene zu entnehmen, auf der die Vergleichsausschnitte zu einer sie vereinigenden Repräsentanz gehören.

Je höher das Repräsentationsniveau, auf dem die für den Vergleich relevante Information entnommen werden kann, desto größer ist auch die zeitliche Erstreckung der auf diesem Niveau befindlichen Binnenglieder, d.h. desto länger ist die zeitliche Invarianzperiode, die dieses Binnenglied konstituiert. Bevor auf einem Repräsentationsniveau einem Binnenglied, das gerade erst entsteht, Informationen über die Beziehung der ihm unmittelbar untergeordneten Binnenglieder entnommen werden kann, muß die Melodie erst einmal so weit fortschreiten, bis die dieses Binnenglied konstituierende Invarianz erkennbar wird. Je länger aber die Invarianzperiode, desto länger muß der Hörer abwarten, bis er sie erkennen kann.

Das heißt nicht notwendigerweise, daß der Hörer immer bis zum Ende der Invarianzperiode abwarten muß, um sie erkennen zu können. Er kann auch aufgrund seiner Erwartungen über den Typ der Invarianz, den seine Wahrnehmungsschemata spezifizieren können, bereits aus einer unvollständigen Invarianzperiode Informationen entnehmen, die mehr für den einen oder den anderen Typ von Invarianz sprechen, um so bereits eine Diskrimination zwischen möglichen Alternativen durchzuführen. Je weiter eine sich entwickelnde Invarianzperiode fortgeschritten ist, desto geringer wird die Zahl der Alternativen von möglichen Invarianztypen, die mit der aufgenommenen Information vereinbar sind. Je länger also eine Invarianzperiode ist, desto länger muß abgewartet werden und Informationen über die sich entwickelnde Invarianz gesammelt werden, um mit k o n s t a n t e r S i c h e r h e i t zwischen alternativen Invarianztypen diskriminieren zu können.

Unsere allgemeine Hypothese lautet also: Je höher das Repräsentationsniveau, das die zu vergleichenden Ausschnitte aus einer Melodie in einer Repräsentanz vereinigt, desto höher ist die für den Vergleich benötigte Zeit.

Wir können den Hörern drei verschiedene Typen von Vergleichsaufgaben stellen:

1. Die Vergleichsausschnitte folgen zeitlich unmittelbar aufeinander (Bedingung "zeitlich nah"), und die Vergleichsausschnitte gehören auf dem unmittelbar übergeordneten Niveau zur selben Repräsentanz (Bedingung "syntaktisch nah").

2. Die Vergleichsausschnitte folgen zwar zeitlich unmittelbar aufeinander (Bedingung "zeitlich nah"), aber sie gehören auf dem unmittelbar übergeordneten Niveau zu zwei verschiedenen Repräsentanzen, die selbst wieder erst auf dem ihnen übergeordneten Niveau zu einer sie vereinenden Repräsentanz gehören (Bedingung: "syntaktisch fern").

3. Die Vergleichsausschnitte werden durch einen Ausschnitt gleicher zeitlicher Erstreckung, wie die der Vergleichsausschnitte selbst, voneinander getrennt (Bedingung "zeitlich fern") und gehören deshalb auf dem ihnen übergeordneten Niveau zu zwei verschiedenen Repräsentanzen, die erst auf dem ihnen selbst wieder übergeordneten Niveau zu einer gemeinsamen Repräsentanz gehören (Bedingung "syntaktisch fern"). (Die Bedingungskombination "syntaktisch nah" und "zeitlich fern" ist nicht realisierbar).

Nach unserer Hypothese sollte die für einen Vergleich benötigte Verarbeitungszeit von der Bedingungskombination "zeitlich und

syntaktisch nah" über die Kombination "zeitlich nah, syntaktisch fern" zur Bedingungskombination "zeitlich und syntaktisch fern" ansteigen. Wählen wir unabhängig von den Bedingungen zeitlicher und syntaktischer Nähe unterschiedlich große Vergleichsausschnitte, z.B. Subphrasen, Takte und Füße, dann ist aufgrund unserer Hypothese zu erwarten, daß die Verarbeitungszeit um so größer ist, je größer die zeitliche Erstreckung der zu vergleichenden Ausschnitte ist.

2.4.2 METHODE

Reizmaterial und Darbietungsweise

Wir boten den Vpn Melodien mit einer achttaktigen Periode dar, die mit Hilfe der Syntax generiert wurden. Die einzige Einschränkung, die bei der Generierung zu berücksichtigen war, bestand darin, daß die Melodien keine Transformationen enthalten sollten bis auf die Wiederholungstransformation zwischen syntaktisch nahen Vergleichsausschnitten. Außerdem mußte zwischen den syntaktisch fernen und ebenso den zeitlich fernen Ausschnitten in der Hälfte der Fälle eine Wiederholung bestehen. Bei zeitlich fernen Vergleichsausschnitten eines Vergleichsniveaus betrug der Abstand zwischen ihnen genau eine Einheit, die dieselbe Größe wie die der Vergleichseinheiten besitzt.

Die Tondauer in den Melodien betrug 25o msec und das ISI 2oo msec. Betonungsmuster, Lautstärke und Frequenzen der Töne wurden wie in Experiment 3 gewählt. Die Melodien wurden binaural über Kopfhörer

dargeboten und von einem Tongenerator erzeugt, der von einem Prozeßrechner gesteuert wurde (vgl. Experiment 1).

Versuchsplan

Der Versuchsplan ist dreifaktoriell mit Meßwiederholungen auf allen drei Faktoren. Die drei Bedingungen syntaktischer und zeitlicher Nähe der Vergleichsausschnitte wurden mit den drei Vergleichsniveaus "Subphrasen", "Takte" und "Füße" komplett kombiniert sowie mit einer same- und einer different-Bedingung, wobei in der same-Bedingung eine Wiederholung zwischen den Vergleichsausschnitten realisiert war, während in der different-Bedingung zwischen den Vergleichsausschnitten keinerlei Beziehung bestand.

Aus diesem Versuchsplan resultieren 3 (Bedingungen syntaktischer und zeitlicher Nähe) X 3 (Vergleichsniveau) X 2 (same/different-Bedingungen) = 18 Bedingungskombinationen, die jeweils viermal dargeboten wurden. Als abhängige Variable wurde die RT für das Vergleichsurteil erhoben.

Versuchsdurchführung

Ein Versuchsdurchgang wurde von einem 2oo msec dauernden Warnton von 25oo Hz eingeleitet. Dies war für die Vpn das Signal, in ihrem Instruktionsheft, das auf jeder Seite eine Angabe über die zu vergleichenden Ausschnitte enthielt, eine Seite umzublättern und die Vergleichsaufgabe zu lesen. Die Vergleichsaufgabe war als

Frage des Typs "Ist y (2. Ausschnitt) identisch mit x (1. Ausschnitt)?" formuliert (Beispiel: "Ist das Vordersatzanschlußglied identisch mit dem Vordersatzmotiv?"). Zum Lesen des Aufgabentextes standen den Vpn 16 sec zur Verfügung. Am Ende dieses Zeitraumes ertönten zwei Warntöne (2500 Hz, Tondauer jeweils 200 msec, ISI 200 msec) und eine Sekunde später begann die Melodie. Zwischen Melodieende und dem folgenden Warnton lagen 5 sec.

Um die Reaktionszeiten zwischen allen Vergleichsniveaubedingungen vergleichbar zu machen, wich in different-Bedingungen bereits der erste Ton des zweiten Ausschnitts vom entsprechenden Ton des ersten Ausschnitts ab und alle folgenden Töne des zweiten Ausschnitts ebenfalls. Außerdem war die Intervallkonfiguration zwischen den Vergleichsausschnitten in diesen Fällen deutlich verschieden. Theoretisch hätte also in jedem Fall das Erkennen des Unterschieds oder der Übereinstimmung zwischen den jeweils ersten Tönen der Vergleichsausschnitte genügt, um eine richtige Reaktion zu ermöglichen.

Die Vergleichsausschnitte wurden in allen Bedingungen gleichmäßig über die Melodien verteilt, d.h. bei Subphrasenvergleichen wurden gleich viele Vergleiche innerhalb des Vordersatzes wie innerhalb des Nachsatzes durchgeführt, und entsprechendes gilt für die anderen Vergleichsebenen. Für same- und different-Bedingungen wurden jeweils dieselben Ausschnitte, aber bei verschiedenen Melodien verglichen. Alle Melodien waren bis auf die zur Realisierung der jeweiligen Bedingung notwendigen strukturellen Eigenschaften vollkommen verschieden.

Die Instruktion erklärt den Vpn die Vergleichsaufgabe an Beispielen. Die zur Bezeichnung der Vergleichsausschnitte verwandten Begriffe wie "Subphrase", "Vorder- oder Nachsatz", "Fuß" usw. waren den Vpn aus dem Syntaxtraining bereits geläufig. In der Instruktion wurde betont, daß so schnell wie möglich reagiert werden solle, Fehler aber nach Möglichkeit vermieden werden sollten. Außerdem wurde den Vpn gesagt, daß im Falle der Nichtübereinstimmung der Vergleichseinheiten bereits der erste Ton des zweiten Vergleichsausschnitts den Unterschied anzeige, so daß im Prinzip sowohl bei Übereinstimmung als auch bei Nichtübereinstimmung durch das Erkennen der Übereinstimmung bzw. des Unterschieds der jeweils ersten Töne der Vergleichsausschnitte richtig reagiert werden könne.

Zu Beginn des Versuchs wurden 36 Übungsdurchgänge, in denen jede Versuchsbedingung zweimal vorkam, durchgeführt. Anschließend wurden 36 experimentelle Durchgänge und nach einer Pause von 15 Min. weitere 38 Durchgänge durchgeführt, wobei nach der Pause die beiden ersten Durchgänge ebenfalls Übungsdurchgänge waren, damit die Vpn sich wieder auf die Aufgabe einstellen konnten. Welche Durchgänge experimentelle oder Übungsdurchgänge waren, wurde den Vpn nicht mitgeteilt. Die Reihenfolge der Versuchsbedingungen war zufällig verteilt und für alle Vpn gleich. Die RT-Messung begann immer unmittelbar nach dem Ende des ersten Tons der zweiten Vergleichseinheit. Die Versuchsdauer betrug durchschnittlich 1 1/2 Std. Es nahmen immer zwei Vpn gleichzeitig am Versuch teil.

Versuchspersonen

Am Versuch nahmen 38 Vpn teil, die das Syntaxtraining absolviert hatten. Die Vpn wurden für diesen Versuch mit 12 DM bezahlt mit Ausnahme der Psychologiestudenten, die ein Testat für abgeleistete Vpn-Stunden erhielten.

2.4.3 BEFUNDE UND DISKUSSION

Auswertungsmethode

Die RTs für fehlerfreie Reaktionen wurden pro Bedingung und Vp gemittelt und mit einer 3x3x2-Varianzanalyse für Meßwiederholungen auf allen Faktoren ausgewertet. Wir verwenden hier wieder die bereits im Experiment 1 dargestellten konservativen Signifikanzkriterien.

Bedingungen syntaktischer und zeitlicher Nähe

Der Haupteffekt für die Bedingungen syntaktischer und zeitlicher Nähe ist sehr signifikant ($F(2,74) = 58.57$, $p<0.001$, bei df = 1,37 ebenfalls $p<0.001$). Wie Abb. 35 zeigt, stimmen die RT-Differenzen mit unserer Hypothese überein (alle Scheffé-Paarvergleiche zwischen Bedingungen sind sehr signifikant, $p<0.001$): Bei syntaktischer und zeitlicher Nähe ist die RT deutlich geringer als bei syntaktischer Ferne und zeitlicher Nähe und in letzterer Bedingung ist die RT ebenfalls deutlich kürzer als in der Bedingung mit syn-

taktischer und zeitlicher Ferne der Vergleichseinheiten.

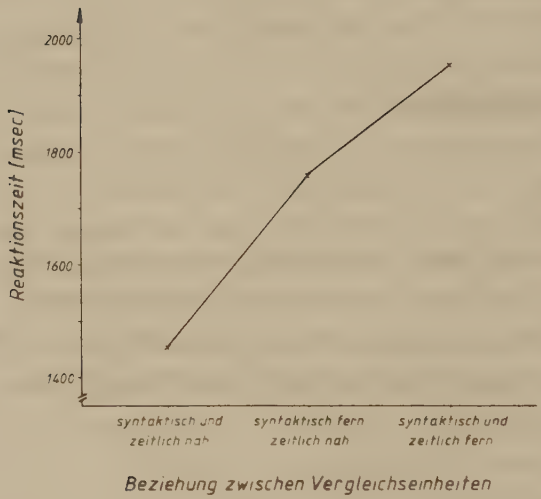


Abb. 35: RT-Mittelwerte in den Bedingungen unterschiedlicher syntaktischer und zeitlicher Nähe

Die Einbettung der Vergleichsausschnitte in die Struktur der übergeordneten Repräsentationsniveaus ist also über die zeitliche Nähe der Vergleichsausschnitte hinaus ausschlaggebend für die Geschwindigkeit, mit der auf die aufgabenrelevante Information zugegriffen werden kann. Nach dem linearen Modell von Massaro (1975, 1977), das von einer zeitlich seriellen Organisation der Tongruppen einer Me-

lodie ausgeht, wäre lediglich ein Effekt der zeitlichen Nähe der Vergleichseinheiten zu erwarten gewesen. Daß sich hier ein ganz deutlicher Effekt der syntaktischen Nähe zeigt, ist ein u.E. überzeugender Beleg für die hierarchische Organisationsstruktur der beim Hören von Melodien gebildeten Repräsentanzen und für die Abhängigkeit der Verarbeitungsprozesse bei der strukturellen Identifikation von der Organisation der Repräsentanzen.

Bedingungen unterschiedlichen Vergleichsniveaus

Der Haupteffekt für das Vergleichsniveau ist signifikant ($F(2,74) = 5.43$, $p<0.01$, bei df = 1,37 $p<0.05$). Abb. 36 zeigt die RT-Mittelwerte für Vergleichsausschnitte auf dem Subphrasen-, Takt- und Fußniveau, die in dieser Reihenfolge abnehmen (Scheffé-Paarvergleiche: Subphrasen/Takte $p<0.05$; Subphrasen/Füße $p<0.01$; Takte/Füße $p>0.10$), wobei allerdings die Abnahme der RT vom Vergleichsniveau der Takte zum Vergleichsniveau der Füße nicht signifikant ist.

Der RT-Abfall zwischen den Vergleichsniveaus "Subphrasen" und "Takte" kann also auf die größere zeitliche Erstreckung der Subphrasen gegenüber den Takten zurückgeführt werden und auf die hierdurch bedingte längere Fokussierung der Variationsstrukturen auf Subphrasenniveau, die zur Konstanthaltung des Sicherheitskriteriums für die Vergleichsurteile notwendig ist. Hätten die Vpn die Möglichkeit genutzt, bereits nach dem ersten Ton des zweiten Vergleichsausschnitts zu reagieren, dann hätten sie bei

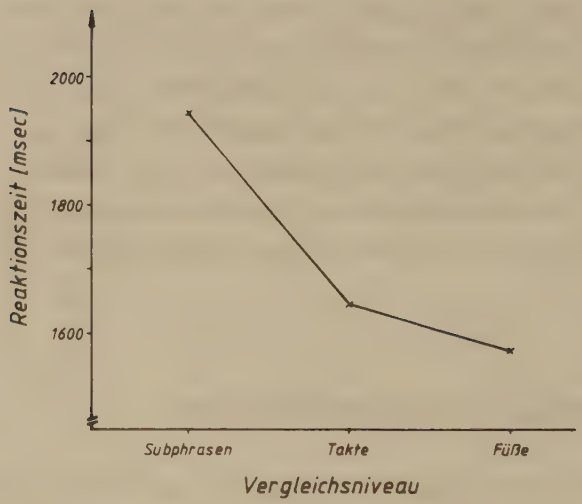


Abb. 36: RT-Mittelwerte in Abhängigkeit vom Vergleichsniveau

Subphrasen, die eine doppelt so große zeitliche Erstreckung besitzen und auch aus doppelt so vielen Tönen bestehen, nur die Hälfte der Information relativ zur Länge des Ausschnitts berücksichtigt, die sie bei Vergleichen zwischen Takten berücksichtigen würden. Aufgrund des RT-Unterschieds zwischen Subphrasen- und Taktvergleichen von 3o2 msec ist anzunehmen, daß die Vpn bei Subphrasen ungefähr einen Ton länger abgewartet haben als bei Taktvergleichen.

Daß die RT bei Vergleichsausschnitten auf dem Niveau der Takte genau so lang ist wie bei Ausschnitten auf dem Niveau der Füße, ist vielleicht dadurch zu erklären, daß die Schwierigkeit der Entnahme aufgabenrelevanter Information über zeitlich sehr kurze Ausschnitte

größer als bei Ausschnitten längerer zeitlicher Erstreckung ist. Ähnliche Schwierigkeiten bei Wahrnehmungsprozessen, die speziell die niedrigsten Repräsentationsniveaus betrafen, haben wir ja auch in den Experimenten 3 und 4 beobachten können (wir werden auf diesen Punkt in der allgemeinen Diskussion noch zurückkommen).

Der Grund für diese Schwierigkeiten könnte darin bestehen, daß sich bei kurzen Ausschnitten die Entscheidung über Gleichheit oder Ungleichheit auf sehr wenig Information stützen muß, die zwar relativ zur Länge der Ausschnitte gleich groß ist, absolut gesehen aber mit niedriger werdendem Vergleichsniveau abnimmt. Der Hörer könnte diesen Schwierigkeiten dadurch begegnen, daß er nach Fokussierung des die Vergleichseinheiten zu einer Repräsentanz integrierenden Repräsentationsniveaus zusätzlich die aufgabenrelevanten Füße fokussiert, um eine Überprüfung seines vorläufigen Vergleichsurteils vorzunehmen. Da ein solcher Verifikationsprozeß nach Bamber (1969) für different-Urteile ohnehin durchgeführt werden sollte, müßte für different-Urteile die RT von den Takten zu den Füßen weiterhin abnehmen, für same-RTs aber größer werden oder zumindest bedeutend geringer als für different-Urteile abnehmen.

Daß diese Interpretation zutreffend sein könnte, zeigen die RT-Mittelwerte der "same/different X Vergleichsniveau"-Bedingungskombinationen in Abb. 37. Die Interaktion ist signifikant ($F(2,74) = 4.07$, $p<0.02$, bei $df = 1,37$ $p<0.10$) und der Steigungsunterschied zwischen same- und different-Bedingungen bei Takt- und Fußvergleichen ebenfalls (Scheffé-Kontrast, $p<0.05$). Die

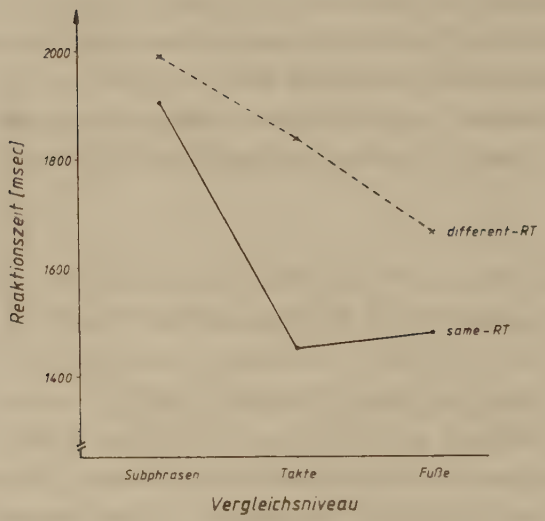


Abb. 37: RT-Mittelwerte für die Interaktion "Vergleichsniveau X same/different"

Analyse der Fehlerhäufigkeiten für die drei Vergleichsbedingungen unterstützt diese Interpretation zumindest der Tendenz nach: Bei Subphrasen und Takten ist die Fehlerzahl gleich hoch (1o.5 % bzw. 9.7 %), bei Füßen ist sie aber geringfügig höher (14.9 %) (Friedman-Test: Chi-Quadrat = 6.73, df = 2, $p \leq o.1o$).

Effekt der same- und different-Bedingungen

Der Haupteffekt für same- und different-Bedingungen ist sehr signifikant ($F(1,37) = 15.84$, $p \leq o.ool$): Same-Vergleiche werden mit einer durchschnittlichen RT von 1611 msec durchgeführt; different-Ver-

gleiche sind demgegenüber bedeutend langsamer, hier liegt die RT bei durchschnittlich 1831 msec.

In Anlehnung an Bamber (1969) sehen wir in diesem Effekt einen Hinweis auf funktionale Unterschiede zwischen same- und different-Vergleichen: Die aufgabenrelevante Information wird bei same-Vergleichen ausschließlich (wenn man einmal vom Fußniveau als Sonderfall absieht) der Repräsentanz entnommen, die die Vergleichseinheiten auf einem übergeordneten Niveau integriert; different-Vergleiche stützen sich zusätzlich noch auf einen Verifikationsprozeß, der in einer Fokussierung der zu vergleichenden Ausschnitte besteht und wegen der hierzu notwendigen Repräsentationsebenenwechsel und zusätzlichen Analyse längere Zeit in Anspruch nimmt als ein same-Vergleich.

Dieser Verifikationsprozeß führt aber offenbar nicht zu einer Reduktion der Fehlerzahlen bei different- gegenüber same-Vergleichen (same: 1o.9 %; different: 12.5 %; Wilcoxon-Test: $z = 1.o4$, $p \geq o.2o$ zweiseitig getestet).

2.5 ZUSAMMENFASSUNG DER BEFUNDE AUS DEN EXPERIMENTEN 3 BIS 5

Die Experimente 3 bis 5 dienten einer Überprüfung der Hypothese, daß die Struktur der Repräsentation beim Hören von Musik hierarchisch ist. Im Sinne konvergierender Operationen versuchten wir mit drei verschiedenen experimentellen Methoden, die einmal stärker den strukturellen Aspekt der Repräsentation (Experiment 3 und 4), das andere Mal stärker den funktionalen Aspekt in den

Mittelpunkt stellten (Experiment 5), nachzuweisen, daß die Periodenstruktur, die die syntaktische Beschreibung der Melodien liefert, mit der Struktur der Repräsentation beim Hören der Melodien durch syntaktisch kompetente Hörer übereinstimmt.

Die Befunde von Experiment 3 unterstützen durch die Verteilung der Klicklokalisationshäufigkeiten diese Hypothese, allerdings mit der Einschränkung, daß die Lokalisationsdaten keine Differenzierung zwischen Ton- und Fußniveau zeigen. Ebenso stützen die Klickdetektionszeiten in Experiment 4 diese Hypothese, hier allerdings mit der Einschränkung, daß die Detektionszeiten nicht zwischen Fuß- und Taktniveau differenzieren.

Zusätzlich wurde in Experiment 4 eine Hypothese über den Unterschied zwischen der kognitiven Verarbeitung von Melodiestrukturen bei syntaktisch trainierten und untrainierten Vpn überprüft. Diese Hypothese besagte, daß syntaktisch trainierte Vpn aufgrund ihrer differenzierteren Wahrnehmungsschemata insbesondere Klicks an den ausgezeichneten Positionen einer Melodie, nämlich an den höheren Konstituentengrenzen, schneller als untrainierte erkennen können sollten, da sie diese Positionen ziemlich genau antizipieren können. Der erwartete Effekt, nämlich kürzere RT bei Klicks an den höheren Konstituentengrenzen bei trainierten Vpn als bei untrainierten, stellte sich zwar ein, erreichte aber für das konservative Signifikanzkriterium nicht mehr die Signifikanzschwelle.

Die sich abzeichnende Tendenz entsprach zwar den Erwartungen, aber nur für die zweite Hälfte der Melodien. Während in der

ersten Hälfte bis kurz vor der Hauptkonstituentengrenze keinerlei Unterschiede im Verlauf der RT-Kurven in Abhängigkeit von der Klickposition zu beobachten waren (wenn man einmal von der stets höheren RT der untrainierten Vpn absieht), zeigte sich im Bereich von der Hauptkonstituentengrenze bis zum Ende der Melodien, daß die untrainierten Vpn offenbar die "Orientierung" verloren hatten: Ihre RTs differenzierten nicht mehr, wie noch am Anfang der Melodien, zwischen der Höhe der Konstituentengrenze, an der sich der Klick jeweils befand. Wir sahen hierin - wenn auch eine sehr schwache - Unterstützung unserer Hypothese.

Mit Experiment 5 gelang uns u.E. der überzeugendste Beleg für unsere Hypothese einer hierarchischen Repräsentationsstruktur. Wir konnten demonstrieren, daß die syntaktische Nähe zweier miteinander zu vergleichender Ausschnitte einer Melodie (operationalisiert durch den vertikalen Abstand zwischen dem Strukturniveau der Vergleichsausschnitte und demjenigen übergeordneten Strukturniveau, auf dem die Vergleichsausschnitte zur selben strukturellen Einheit gehören) über die zeitliche Nähe der Ausschnitte hinaus die für den Vergleich benötigte Zeit determiniert: Die RT für zeitlich benachbarte Ausschnitte war für gleichzeitig syntaktisch nahe Ausschnitte sehr signifikant niedriger als für syntaktisch ferne Ausschnitte und die für syntaktisch und gleichzeitig zeitlich ferne Ausschnitte war noch bedeutend höher als die für letztere Ausschnitte.

In der folgenden Diskussion einiger Aspekte unserer Befunde aus den Experimenten 1 bis 5 werden wir u.a. die kognitive Adäquatheit der Syntax, die den in den Experimenten dargebotenen Melodien zugrundelag, zu beurteilen und den theoretischen Rahmen des Hierarchiemodells auf Höreinstellungen zu übertragen versuchen, die nicht so offensichtlich von kognitiven Prozessen bestimmt sind, wie das strukturelle Hören, das im Mittelpunkt unserer Untersuchung stand.

3. ALLGEMEINE DISKUSSION

Ernst Kurth's Musikpsychologie

Die unmittelbaren Anregungen zu dieser Untersuchung, insbesondere für die in Teil II thematisierte Fragestellung, erhielten wir von der modernen Linguistik und Psycholinguistik. Aber bereits Kurth (1931) hat den von uns eingeschlagenen Weg vorgezeichnet: "Die Musikpsychologie muß zur Theorie greifen, ähnlich wie man aus der Grammatik eine Psychologie der Sprache rekonstruieren kann" (Kurth, 1931, S. 65). Gemeint ist hier, daß die Musiktheorie für die Musikpsychologie dieselbe Funktion übernehmen müsse, die die Grammatiktheorie für eine Psychologie der Sprache besitzt: Sie sollte die Strukturen von Musik in einer Weise aufklären, die es dem Musikpsychologen gestattet, Hypothesen über die Wahrnehmung dieser Strukturen zu formulieren; und die musikpsychologischen Befunde sollten wiederum Eingang in die Musiktheorie finden, damit die Form ihrer Beschreibung der musikalischen Strukturen den psychischen Bedingungen ihrer Wahrnehmung gerecht wird.

Kurth wendet sich gegen die elementaristische Betrachtung des Psychischen, die den einzelnen Ton als "Tonvorstellung" (Riemann, 1914/15) oder die Wahrnehmung von Beziehungen zwischen Tönen als klassisch konditioniert (Schole, 193o) bezeichnet. Er vertritt die gestalttheoretische Auffassung, daß "das Melodische keine Zusammenfassung von Tönen ist, sondern ein ursprünglicher Zusammenhang, aus dem sich Töne herauslösen" (Kurth, 1931, S. 78).

"Gestaltbewegungen" oder "Einheitsbilder", wie Kurth die psychische

Repräsentation wahrgenommener Musik nennt, entwickeln sich im Verlauf der Wahrnehmung vom einzelnen Ton zu Einheiten immer höherer Ordnung: "Im Fortschreiten zu immer größeren Phänomenen erkennt man stetig die neue Überwölbung von Teilerscheinungen durch einen Einheitseindruck, in dem deren Vielfalt zurückgedrängt und die Komplexitätsempfindung vorgetrieben wird" (Kurth, 1931, S. 33). Diese Sicht des Wahrnehmungsvorgangs ist gar nicht so unähnlich unserer These, daß die Repräsentation von Musik eine hierarchische Struktur besitzt, die bei Auftreten von Invarianzen höherer Ordnung und längerer zeitlicher Periode durch die Bildung immer höherer Repräsentationsniveaus während des Hörens allmählich aufgebaut wird.

Durch den Hinweis auf diese Parallelen wollen wir nicht für eine Renaissance der Gestalttheorie plädieren, vielmehr meinen wir, daß es nach jahrzehntelanger "Opposition" des Informationsverarbeitungsansatzes gegenüber gestalttheoretischen Einsichten an der Zeit ist, diese erneut als Anregungen aufzugreifen und in die theoretischen Überlegungen einzubeziehen, wie dies z.B. Bregman (1978a) bereits tut. Hierdurch könnte man unserer Ansicht nach von der Eingleisigkeit der Computeranalogie wegkommen und zu einer Revision der Modellannahmen des linearen Informationsverarbeitungsansatzes gelangen, wie z.B. der Linearitätsannahme über die Prozeßabfolge und der noch überwiegend elementaristischen Sichtweise der Wahrnehmungsreize.

Kognitiv adäquate Aspekte der Syntax

In unseren Experimenten konnten wir nur den kleineren Teil von Hypo-

thesen über die Wahrnehmung von Musik überprüfen, die sich bei der Übersetzung des musikalisch-linguistischen Syntaxmodells in eine psychologische Theorie der Musikwahrnehmung ergeben. Im Mittelpunkt unserer Untersuchung stand die Hypothese einer strukturellen Übereinstimmung zwischen der Repräsentation der Melodien beim Hören und den bei der linguistischen Analyse dieser Melodien aufzudeckenden Beziehungsstrukturen zwischen ihren Einheiten verschiedener Ordnung. Diese Hypothese hat sich im Verlauf der Untersuchung als sehr brauchbar erwiesen: Fehllokalisationen von Klicks (Experiment 3), Detektionszeiten für Klicks (Experiment 4) und Reaktionszeiten für Vergleiche zwischen syntaktisch definierten Ausschnitten von Melodien (Experiment 5) sind mit der Hypothese einer hierarchischen Repräsentationsstruktur gut vereinbar.

Die Reichweite der Verallgemeinerbarkeit dieser Hypothese ist jedoch noch stark eingeschränkt, da wir in unseren Experimenten lediglich regelmäßige, achttaktige Periodenstrukturen verwendet haben. Berücksichtigt wurden keine Perioden mit Mittelsatz, keine Erweiterungen der Periode durch Motivsequenz im Vordersatz oder Anschlußgliedwiederholung im Vorder- und Nachsatz, keine Erweiterungen von Anschlußgliedern um einen Takt und ebensowenig die möglichen Formen der Periodenkürzungen. Bereits die strukturell einfachste musikalische Gattung, die der Volks- und Kinderlieder, ist so variantenreich, daß eine Untersuchung, die der vollen Komplexität solcher musikalischen Strukturen gerecht werden könnte, in absehbarer Zeit nicht zu erwarten ist. Trotz dieser Einschränkungen sind wir zuversichtlich, daß sich die an exemplarischen Melodiestrukturen nachzuweisende hierarchische Form der Repräsentation

auch bei Verwendung weniger regelmäßig aufgebauter musikalischer Perioden hätte demonstrieren lassen.

Die syntaktische Struktur der musikalischen Einheiten und ihrer Beziehung zur Struktur der kognitiven Repräsentation stellt jedoch nur einen Aspekt der kognitiven Adäquatheit einer musikalischen Syntax dar. Ein weiterer wichtiger Aspekt ist die Form der verwendeten Regeln. Die beiden wichtigsten Regeltypen, die in der syntaktischen Beschreibung vorkommen, sind die Ersetzungsregeln und die ordinalen Transformationsregeln.

Chomsky (1964, Anm. 16) warnt vor einer allzu leichtfertigen Übertragung der Regeln einer generativen Grammatik auf die Funktionsweise kognitiver Prozesse, aber diese Warnung ist längere Zeit ungehört geblieben (vgl. McNeill, 197o; Miller & Isard, 1967; Saporta, 1967), und bei Chomsky (1968) gewinnt man zuweilen den Eindruck, als ob er sie selbst nicht immer beachten würde. Wenn wir im folgenden die Form der syntaktischen Regeln mit kognitiven Prozessen in Beziehung bringen, dann meinen wir nicht die k o n s t r u k t i v e R o l l e, die diese Regeln für die Ableitung eines musikalischen Kontextes spielen, sondern alleine ihre g e n e r a l i s i e r e n d e R o l l e, die in der Formulierung von Regelmäßigkeiten innerhalb eines Kontextes bestehen (Olshowsky, 1975).

In dieser Hinsicht beschreiben die Ersetzungsregeln des Typs X→YZ denselben Sachverhalt, den das Hierarchiemodell als "Binnengliederung einer Repräsentanz X durch die ihr untergeordneten Repräsentanzen Y und Z" beschreibt, also die horizontale Gliede-

rung auf einem Repräsentationsniveau beim Übergang von einem höheren auf ein niedrigeres Repräsentationsniveau. Die Struktur, die diese Regeln beschreiben, ist eine hierarchische, und insofern, als die Struktur der kognitiven Repräsentation dieselbe Form der Gliederung zeigt, sind diese Regeln als kognitiv adäquat zu bezeichnen. Da aber die Regeln unter Berücksichtigung ihrer konstruktiven Rolle die Gliederung von der höchsten strukturellen Ebene zur niedrigsten beschreiben, während der Aufbau der Repräsentationsstruktur analog der zeitlichen Entwicklung einer Melodie von der niedrigsten zur höchsten Ebene fortschreitet, müßten die Regeln eigentlich umgekehrt formuliert werden, d.h. als YZ⟵X geschrieben werden, womit sie nicht mehr Regeln einer generativen, sondern einer analytischen Syntax wären.

Über die kognitive Adäquatheit der ordinalen Transformationsregeln, die im Unterschied zu den Transformationsregeln der Generativen Grammatik Chomsky's (1965) nicht nur auf eine strukturelle Ebene (die Basis) angewendet werden können, sondern auf alle, und die nicht eine erst vollständig vorliegende Tiefenstruktur in eine Oberflächenstruktur überführen, sondern einen vorausgehenden Kontext eines beliebigen strukturellen Niveaus in einen nachfolgenden Kontext desselben Niveaus überführen, können wir aufgrund unserer Untersuchungen leider noch keine empirisch fundierten Aussagen machen.

Das einzige Experiment, in dem solche Transformationen eine systematische Rolle spielten, war das Experiment 5. Wir haben diese Transformationen mit dem Invarianzkonzept in Beziehung gebracht, und zwar mit dem Konzept der transformationellen Invarianzen, die

den Stil einer sich über die Zeit erstreckenden systematischen Veränderung charakterisieren (Shaw & Pittenger, 1978). In Experiment 5 nahmen wir an, daß das Erkennen einer solchen Invarianz um so länger dauern sollte, je größer die zeitliche Erstreckung der Invarianzperiode ist. Diese Hypothese konnte durch die Daten von Experiment 5 unterstützt werden. Allerdings benutzten wir lediglich die Transformation "Wiederholung", so daß eine Verallgemeinerung der in Experiment 5 ermittelten Befunde auf das Erkennen anderer Transformationsbeziehungen gegenwärtig noch nicht möglich ist, obwohl wir kaum glauben, daß sich bei Verwendung anderer Transformationsbeziehungen in der in Experiment 5 gestellten Erkennungsaufgabe grundsätzlich andere Befunde einstellen würden.

Leonard Bernstein, der gefeierte Dirigent und Komponist, hat sich in seinen Harvard Vorlesungen (Bernstein, 1979) in das, wie er es nennt, "Chomsky-Land" begeben und in einer auch für den Laien verständlichen Form demonstriert, zu welchen Einsichten eine mit linguistischen Methoden arbeitende musikalische Analysetechnik gelangen kann. Seine Analyse der "Tiefenstruktur" von Mozart's Sinfonie Nr. 4o (KV 55o) ist interessant, obwohl sie oder gerade weil sie auf einem Mißverständnis des Transformationskonzepts von Chomsky gegründet ist. Chomsky's Transformationen werden auf die vollständig vorliegende Basis angewendet bzw. in der Analyse aus der vollständig vorliegenden Oberflächenstruktur erschlossen. Wenn diese Konzeption kognitiv adäquat sein sollte, müßte der Sprecher zuerst das kognitive Äquivalent einer vollständigen Basis des Satzes im Gedächtnis speichern, bevor er Transformationen vornehmen könnte. Kognitiv wesentlich ökono-

mischer wäre hingegen eine gleichzeitige Generierung von Basis und Transformationen, weil die hierfür zu postulierende Gedächtniskapazität bedeutend geringer sein kann, wenn ein bereits generierter Satzteil ausgesprochen werden kann, bevor die Basis der weiteren Satzteile vollständig generiert worden ist. Ohne den Unterschied zum Transformationskonzept der Generativen Grammatik zu bemerken, geht Bernstein davon aus, daß die Transformationen "Tilgung" (Verkürzung einer melodie-syntaktischen Einheit) und "Permutation" (Umordnung der Reihenfolge von Tönen einer syntaktischen Einheit) in der zeitlichen Reihenfolge des Auftretens dieser Teilstrukturen und nicht erst im Anschluß an die vollständige Ableitung der Basis vorgenommen werden können. Dieses Verständnis von Transformationen in Musik ist dem von Ruwet (1972) und Schenker (1954) sehr ähnlich und entspricht genau unserer Intention, die wir bei der Formulierung ordinaler Transformationsregeln verfolgt haben, Transformationen in die syntaktische Beschreibung an den Stellen der Ableitung einzufügen, an denen der Hörer bei akustischer Realisierung dieser musikalischen Strukturen solche Transformationen erkennen kann.

Die Anzahl der von uns bei der syntaktischen Analyse eines Corpus von Volks- und Kinderliedern gefundenen Transformationen ist noch recht bescheiden. Bei der Analyse des Corpus hatten wir einige Male den Eindruck, eine bestimmte Passage müsse sich eigentlich in Form einer ordinalen Transformation darstellen lassen, aber diese Darstellung ist uns dann nicht gelungen. Vielleicht waren hier unsere Wahrnehmungsschemata unserem expliziten Wissen über musikalische Strukturen bereits einen Schritt voraus. Wir meinen, daß die systematische Suche nach ordinalen Transformationsbeziehungen in musikalischen Strukturen eine für die musikalisch-

linguistische Analyse erfolgversprechende Forschungsstrategie darstellen könnte, die zur Erhöhung der kognitiven Adäquatheit musikalisch-syntaktischer Beschreibungen beitragen könnte.

Funktionen musikalischen Wissens für die Wahrnehmung von Musik

Die Befunde von Experiment 4 demonstrieren eine zwar schwache, aber in Richtung unserer Hypothese gehende Tendenz, die Unterschiede in der Wahrnehmung von Melodien zwischen syntaktisch trainierten und untrainierten Hörern offenbaren. Die Frage, die wir hier aufgreifen wollen, ist die nach den Funktionen syntaktischen Wissens für die Wahrnehmung von Melodien, speziell die Frage, ob die Befunde von Experiment 4 bereits eine Funktion musikalisch-syntaktischen Wissens demonstriert haben.

Bevor wir diese Frage beantworten können, müssen wir erst eine etwas differenziertere Vorstellung darüber entwickeln, was eigentlich unter einem musikalisch-syntaktischen Wissen zu verstehen ist.

Chomsky (1965) spricht davon, daß der Sprachbenutzer über ein implizites oder intuitives Wissen über die Regeln verfüge, nach denen die grammatisch wohlgeformten Sätze einer Sprache gebildet werden, und sein Ziel war es, dieses Wissen in Form eines rationalen, explizit formulierten Modells darzustellen. Die Generative Grammatik versteht sich deshalb nicht nur als Modell der Sprachstrukturen, sondern auch als idealisiertes Modell des Wissens über Sprache. Die Suche nach der "psychologischen Realität" der Grammatik (z.B. Miller, 1962; Miller & Isard, 1967) war

u.a. der Frage gewidmet, ob die Form, in der dieses Wissen kognitiv repräsentiert ist, mit der Form, in der die Generative Grammatik dieses Wissen zu beschreiben glaubt, übereinstimmt. Daß diese Fragestellung an der psychologischen Realität vorbeigeht, haben wir bereits in der Einleitung unserer Untersuchung dargelegt (vgl. auch Hörmann, 1976). Uns interessiert hier auch nicht die Frage, ob Inhalt und Form der musikalischen Syntax, die den Melodien unserer Experimente zugrundeliegt, mit Inhalt und Form der Repräsentation eines musikalischen Wissens übereinstimmt, sondern unabhängig davon die Frage, welche Funktionen ein beim Hörer evtl. vorhandenes musikalisches Wissen für die Wahrnehmungsprozesse haben könnte.

Dabei kommen wir aber nicht an der Beantwortung der Frage vorbei, durch welche Repräsentationsform und welchen Inhalt musikalisches Wissen zu charakterisieren ist. Musikalisches Wissen ist, wie auch das über die Strukturen der Sprache, größtenteils ein implizites Wissen, dessen wir uns im Augenblick seiner "Anwendung" nicht bewußt sind:

"We perform a skill by relying on the coordination of elementary ... acts, and we are aware of having got these right by accomplishing our skillful performance. We are aware of them in terms of this performance and not (or very incompletely) aware of them in themselves... There are vast domains of knowledge... that exemplify... that we are generally unable to tell what particulars we are aware of when attending to a coherent entity which they constitute. Thus there are two kinds of knowing which invariably enter jointly into any act of knowing a comprehensive entity. There is (1) knowing a thing by attending to it, in the way we attend to an entity as a whole and (2) knowing a thing by relying on our awareness of it for the purpose of attending to an entity to which it contributes" (Polanyi, 1962, S. 601).

Ersteres Wissen, von dem Polanyi (1962) hier spricht, ist ein "fokales" oder "explizites" Wissen, dessen wir uns bewußt sind, wenn wir es anwenden, letzteres dagegen ein "subsidiäres" oder "implizites" Wissen, über dessen Existenz wir erst dadurch etwas erfahren, daß wir eine Handlung, die dieses Wissen voraussetzt, erfolgreich durchgeführt haben.

Genauso wie wir uns unseres Wissens über Sprache während des Sprechens und Hörens nicht ständig bewußt sind, sind wir uns unseres Wissens über musikalische Strukturen nicht bewußt, wenn wir Musik hören oder uns reproduktiv musikalisch betätigen. Die allermeisten Musikhörer können dieses Wissen auch nicht durch sprachliche Bezeichnungen explizit machen, und wenn ein Hörer dies kann, tut er es während des Hörens nur in den seltensten Fällen.

Die Form, in der solches Wissen kognitiv repräsentiert ist, muß wohl in dem Sinne sehr abstrakter Natur sein, daß es nicht konkrete musikalische Ereignisse, sondern vielmehr kategoriale Eigenschaften musikalischer Strukturen sind, die repräsentiert sein müssen. Adorno (1956) vergleicht musikalische Kategorien mit "primitiven Begriffen":

"Man pflegt das Unterscheidende (zwischen Sprache und Musik, Anm. des Verf.) darin zu suchen, daß Musik den Begriff nicht kenne. Aber manches in ihr kommt den 'primitiven Begriffen' recht nahe, von denen die Erkenntnistheorie handelt. Sie benutzt wiederkehrende Sigel. Geprägt wurden sie von der Tonalität. Wenn nicht Begriffe, so zeitigte diese doch Vokabeln: vorab die stets wieder mit identischer Funktion einsetzenden Akkorde, auch eingeschliffene Verbindungen wie die der Kadenzstufen, vielfach selbst melodische Floskeln, welche die Harmonie

umschreiben... Ihre Invarianz hat sich gleichwie eine zweite Natur sedimentiert" (Adorno, 1956, S. 72).

Musikalische "Begriffe" sind i n v a r i a n t e K a t e g o r i e n , die sich als "Sediment" des bei aller Variation zwischen musikalischen Kontexten Konstanten herausbilden. Sie sind in einem zweiten Sinne daher durchaus auch als konkrete Repräsentationen zu betrachten, nämlich als Invarianzkategorien mit einem "reiznahen" Bezug zur Wahrnehmung. Reber (1967, 1969, 1976; Reber & Lewis, 1977) sieht den Prozeß des Erwerbs solcher Invarianzkategorien als ein Wahrnehmungslernen im Sinne von Gibson & Gibson (1955), in dessen Verlauf eine Differenzierung zwischen strukturell unterschiedlichen Reizmustern gelernt und das ihnen strukturell Gemeinsame als implizites, syntaktisches Wissen repräsentiert wird.

Ob man das Format der Repräsentation solcher Invarianzkategorien als Regelstruktur (wie Reber, 1967) oder als Prototyp (wie Rosch, 1978) betrachten kann, ist keineswegs entschieden. Wir verwenden hier deshalb weiterhin den inhaltlich weniger spezifischen Begriff des Schemas (Neisser, 1976) als Repräsentation musikalischer Invarianzkategorien.

Die Tonalität ist als eine über den individuellen Einzelfall eines Musikstücks hinausgehende Form der Invarianz zu betrachten; sie ist strukturelles Bezugssystem der Töne und Tongruppen innerhalb wie zwischen Musikstücken. Wir können sie deshalb als die wahrscheinlich globalste Form musikalischer Invarianzen bezeichnen, die Inhalt oder Spezifikation des höchsten Wahrnehmungsschemas eines nach Invarianzordnungen hierarchisch organisierten Systems

von Wahrnehmungsschemata ist.

Die Bezugssystemfunktion dieses Schemas drückt sich darin aus, daß es einem Hörer, der keinerlei Erfahrung mit atonaler Musik besitzt, nahezu unmöglich ist, in atonaler Musik irgendwelche Strukturen zu erkennen. Die Invarianz der Tonalität in der Musik bis zum Ende des 19. Jahrhunderts mache es dem Bewußtsein so schwer, von ihr Abschied zu nehmen, meint Adorno (1956), aber leider nur deshalb, so müssen wir ergänzen, weil es mangels Wahrnehmungserfahrungen mit atonaler Musik dem kognitiven System unmöglich ist, ein neues globales Bezugssystem hierfür aufzubauen; doch solange dieses nicht in Ansätzen zumindest vorliegt, wird die neue Hörerfahrung zu vermeiden versucht, so daß ein das tonale Bezugssystem stabilisierender circulus vitiosus einsetzt.

Des Bezugssystems beraubt begegnet der Hörer atonaler Musik notwendigerweise zuerst mit Unverständnis, denn wenn die Hypothese einer hierarchischen Verschachtelung der Wahrnehmungsschemata, an deren Spitze das der Tonalität steht, eine adäquate Beschreibung der kognitiven Organisation musikalischen Wissens darstellt, dann sollten selbst vertraute Invarianztypen niedrigerer Ordnung innerhalb des atonalen Kontextes schwerer zu erkennen sein als innerhalb des vertrauten Bezugssystems (vgl. auch Werner, 1926, 194o). Selbst so einfache Intervallkonfigurationen, wie z.B. eine aufsteigende Tonfolge, wird bei Variation der Intervallgrößen schlechter erkannt als bei relativ konstanten Intervallgrößen (vgl. Experiment 1 und 2). Die Ungewöhnlichkeit der Intervallprogressionen in atonaler Musik für einen ganz an der Tonalität orientierten Hörer dürfte die Resonanz mit einer Invarianzkategorie beträchtlich erschweren.

Unabhängig von dem andersartigen Bezugssystem macht es die serielle Musik dem Hörer dadurch besonders schwer, Strukturen zu erkennen, weil bereits die strukturell niedrigsten Transformationsbeziehungen eine zeitliche Periode besitzen, die weitaus größer ist, als die der kleinsten Transformationsbeziehung in tonaler Musik. Das Erkennen z.B. eines Krebses in einer Zwölftonkomposition erfordert vom Hörer die Berücksichtigung eines 24 Töne zählenden musikalischen Ausschnitts. Dagegen kann in klassisch-tonaler Musik die kürzeste Invarianzperiode sich auf einen einzigen Takt beschränken (z.B. Sequenz des zweiten gegenüber dem ersten Fuß).

Höhere Formen der Invarianz als Invarianz der Beziehungen zwischen Ausschnitten niedrigerer Ordnung, die zu einer sehr regelmäßig gegliederten hierarchischen Repräsentationsstruktur führen, fehlen in Zwölftonmusik so gut wie vollständig, da die Transformationen überwiegend seriell angelegt sind. Dies muß notwendigerweise zu einer kognitiven Repräsentation führen, die über weniger Repräsentationsniveaus als bei klassisch-tonaler Musik verfügt, aber gleichzeitig zeitlich bedeutend weiter gespannte Invarianzperioden aufweist.

Das an der Tonalität orientierte und für hierarchisch vielschichtig strukturierte Repräsentanzen adaptierte System von Wahrnehmungsschemata ist dann zur aktiven Steuerung der Wahrnehmungsprozesse durch Antizipation strukturell markierter Nahtstellen und Bereitstellung differenzierter kognitiver Schemata nicht fähig und der Hörer somit zu einem passiven, lediglich die Variationsstrukturen der Merkmale verfolgenden Hören verurteilt. Daß dies Konsequenzen für das affektive Erleben und die ästhetische Stellungnahme des Hörers haben sollte, dürfte auf der Hand liegen (einige experimentelle Befunde zur ästhetischen Beurteilung und zu affektiven Reaktionen auf Tonfolgen lassen sich hierzu als Beleg anführen, vgl. Vitz, 1974).

Ebenfalls als Invarianzen eines sehr globalen Niveaus, aber bereits geringerer Ordnung als die der Tonalität, können musikalische Stile betrachtet werden, die an der Tonalität orientiert sind. Nahezu jeder Hörer kann innerhalb des von ihm favorisierten Bereichs von Musik stilistische Unterschiede erkennen, während er dies in einigen anderen Bereichen nicht zu tun vermag. Diese Abhängigkeit von der individuellen Hörerfahrung zeigt uns, daß hierbei implizite Wissensstrukturen eine bedeutende Rolle spielen müssen. Das stilistische Wissen musikalisch geschulter Hörer kann so differenziert sein, daß sie in der Lage sind, z.B. zwischen unbekannten Volksliedern und Melodien graduell fein abgestuft zu diskriminieren, die in unterschiedlichem Ausmaß Ab-

weichungen von typischen Volksliedstrukturen aufweisen (Stoffer, 1979).

Eine relativ zu diesen globalen musikalischen Invarianzen niedrige Invarianzordnung besitzt die Invarianz der Akzentverteilung, die einen Rhythmus konstituiert. Daß bei der Wahrnehmung von Rhythmen ein musikalisches Wissen eine Rolle spielen sollte, ist vielleicht am wenigsten unmittelbar einsichtig, weil der Rhythmus ein so dominantes musikalisches Merkmal darstellt, das jedermann zu identifizieren in der Lage ist, daß ein Einfluß kognitiver Strukturen in diesem Fall unwahrscheinlich erscheinen muß. Und doch zeigt das Beispiel polyrhythmischer Strukturen, die der abendländischen Musiktradition bis zum Ende des 19. Jahrhunderts fremd waren, daß offenbar auch für die Identifikation von Rhythmen differenzierte Wahrnehmungsschemata notwendig sind.

Der Rhythmus besitzt aber für die Verarbeitung von Musik und auch von Sprache eine besondere Funktion: "... since rhythmically patterned sounds have a time trajectory that can be tracked without continuous monitoring, perception of initial elements in a pattern allows later elements to be anticipated in real time" (Martin, 1972, S. 488). Die Ausnutzung rhythmischer Informationen eines Musikstücks erlaubt es dem Hörer, die kontinuierliche Variationsstruktur diskontinuierlich zu verfolgen, d.h. sie zu segmentieren, um die Segmente einer Weiterverarbeitung, z.B. einer Identifikation, zuzuführen und/oder ihr antizipatorisch vorauszueilen, um Wahrnehmungsschemata zu aktivieren, die für eine Informationsentnahme über zukünftige musikalische Teilstrukturen geeignet sind, oder eine Reaktionsprogrammierung einzuleiten, die dann nur noch durch aktuelle

Informationen gesteuert zu werden braucht. Eine solche antizipatorische Reaktionsprogrammierung findet wahrscheinlich beim Spielen eines Musikinstruments statt, um die zeitliche Koordination der Fingerbewegungen über größere Zeitabschnitte hinweg planen zu können.

Antizipationen sind aber nur dann möglich, wenn das kognitive System die einen Rhythmus charakterisierende Periode erkannt hat und, obwohl erst eine solche Periode zu Beginn eines Musikstücks wahrgenommen worden sein kann, von Anfang an von der Hypothese ihrer Konstanz ausgeht. Die Unterschiede zwischen der syntaktisch trainierten experimentellen Gruppe und der Kontrollgruppe in Experiment 4 bezüglich der Geschwindigkeit der Klickdetektion haben wir durch einen solchen Antizipationsprozeß zu erklären versucht, der es den Hörern der experimentellen Gruppe eher als denen der Kontrollgruppe erlaubt, sich auf das Entdecken eines Klicks an den ausgezeichneten Stellen innerhalb der Periodenstruktur einzustellen. Da diese Stellen mit Taktgrenzen und damit genau mit einer rhythmischen Periode übereinstimmen, ist anzunehmen, daß der Rhythmus der Melodien die strukturelle Grundlage für die antizipatorischen Prozesse der Hörer in der experimentellen Gruppe darstellte.

Wenn sich die Erwartungsbildung der Hörer in der experimentellen Gruppe auf mehrere Ebenen der Repräsentationshierarchie bezieht, dann sollten diese bereits bei Beginn der ersten rhythmischen Periode nicht nur deren Ende, sondern auch bereits das des Periodenpaares antizipieren können, wo die übergeordnete syntaktische Grenze zwischen Motiv und Anschlußglied liegt. Da die Hörer der Kon-

trollgruppe aber über keine derart spezifischen Kenntnisse hinsichtlich der Regelmäßigkeiten der Periodenstruktur verfügen, die es ihnen gestattet hätten, vergleichbar weitreichende Antizipationen struktureller Grenzen vorzunehmen, haben diese wahrscheinlich bei Zunahme der strukturellen Komplexität der Melodien durch das Auftreten mehrerer übergeordneter Invarianzen in der zweiten Melodiehälfte das aktive Antizipieren der jeweils nächstmöglichen Klickposition eingestellt, denn ziemlich genau von der Mitte der Melodien an zeigen die Detektionszeiten der Vpn der Kontrollgruppe keinerlei Beziehung mehr zur Struktur der Melodien. Es war also weniger die Unfähigkeit der Vpn der Kontrollgruppe, den Rhythmus der Melodien wahrzunehmen, als ihre Unfähigkeit, die rhythmische Regelmäßigkeit zur Gliederung der im Verlauf der Entwicklung der Melodie komplexer werdenden Melodiestruktur heranzuziehen, die zu dem unterschiedlichen Detektionsverhalten der beiden Vpn-Gruppen beigetragen hat; sie haben sozusagen über die musikalische Struktur die Übersicht verloren und waren deshalb gezwungen, ein relativ niedriges Repräsentationsniveau zu fokussieren, um ihre Aufmerksamkeit auf den jeweils nächsten Tonzwischenraum zu richten, während die Vpn der experimentellen Gruppe auch Informationen über höhere Repräsentationsniveaus in ihre antizipatorischen Reaktionsvorbereitungen einbeziehen konnten.

In dieser Interpretation reflektieren die Unterschiede zwischen beiden Vpn-Gruppen also nicht unterschiedlich strukturierte Repräsentationen der wahrgenommenen Melodien, sondern unterschiedliche Strategien des Zugriffs auf die Repräsentationsstruktur und der Entnahme von Informationen, die der Reaktionsvorbereitung

dienen.

Wenn im Experiment 4 die differenzierteren kognitiven Strukturen der syntaktisch trainierten Vpn erst dazu geführt haben, daß ihr Detektionsverhalten Beziehungen zur syntaktischen Struktur der Melodien zeigt, dann sollten die deutlich von der syntaktischen Struktur abhängigen Lokalisationsurteile in Experiment 3 ebenfalls hierauf zurückzuführen sein. Ist also die Interpretation, die wir aufgrund der Befunde von Experiment 3 gegeben haben, falsch, d.h. reflektiert das Muster der Lokalisationshäufigkeiten gar nicht die Struktur der Repräsentation, sondern vielmehr die von den Hörern aktiv vorgenommenen Strukturierungen? Genau diese alternative Interpretation wurde ja immer für die Befunde der psycholinguistischen Klickexperimente gegeben und zwar deshalb, weil sprachliche Einheiten als solche in der physischen Signalstruktur nicht nachzuweisen seien (Joos, 195o; Sanders, 1977).

Martin (1972) hat dagegen überzeugende Argumente und Beispiele dafür vorgebracht, daß eine definitive Beziehung zwischen grammatischer Struktur und dem akustischen Muster besteht. Seiner Ansicht nach konstituiert das zeitliche Muster des akustischen Signals eine rhythmische Struktur der Sprache, die durch die zeitlichen Begrenzungen des motorischen Teils der Sprachproduktion entstehen, d.h. die motorische Ausführung ist durch die natürlichen Grenzen der Artikulation bestimmt, die sich in der Organisation des akustischen Signals niederschlagen.

"In the case of speech, these constraints influence the segmental details of the acoustic signal, but they should be expected also to affect at least some aspects of the morphology and syntax of any language at the level of syllable strings since they play a

role in determining what can be naturally and easily spoken. The constraints on speech sounds, or on any other real-time sequence of behavioral elements, that is directly implied by the concept of rhythm is r e l a t i v e t i m i n g , which means that the locus of each (sound) element along the time dimension is determined relative to the locus of a l l other elements in the sequence, adjacent a n d non-adjacent. That is to say that sequences of sounds, speech or otherwise, that are rhythmic will possess hierarchical organization, that is, a coherent internal structure, at the sound level" (Martin, 1972, S. 488).

Wenn also die physische Signalstruktur bereits eine hierarchische Organisation aufweist, dann sollte die Wahrnehmung zu einer analogen, hierarchischen Reizmusterrepräsentation führen und die Prozesse des Zugriffs auf die Repräsentationsstruktur sollten hiervon systematisch beeinflußt werden. Martin (1972) nimmt nun an, daß die rhythmische Struktur unabhängig von der Information über die Segmentidentität, die zwischen zwei Betonungen liegt, verarbeitet werden kann. In einer ersten Verarbeitungsstufe werden die rhythmisch definierten Aussschnitte des Signalmusters zu Verarbeitungseinheiten gruppiert. Diese Verarbeitung besitzt in Martin's Modell dieselbe Funktion, die die Fourieranalyse in unserem Modell besitzt. Erst in einem zweiten Verarbeitungsschritt wird die Information auf niedrigeren strukturellen Niveaus, deren Binnengliederung durch die erste Verarbeitungsstufe geleistet wurde, weiterverarbeitet und identifiziert.

Wenn die rhythmische Struktur der akustischen Signalfolge mit der syntaktischen Struktur der Sprache oder der Musik korreliert ist und zwar auf allen Niveaus der rhythmischen Hierarchie, dann braucht die Repräsentation der wahrgenommenen Signalfolge also nicht mehr unter Verwendung kognitiver Schemata nach syntaktischen Gesichtspunkten aktiv segmentiert zu werden, sondern die rhythmisch ausge-

gliederten syntaktischen Einheiten können auf allen Niveaus der Repräsentationshierarchie direkt analysiert werden.

In musikalischen Strukturen ist der Rhythmus mit der syntaktischen Struktur perfekt korreliert: Bis zum Taktniveau herab fallen alle syntaktischen Grenzen innerhalb der Periodenstruktur mit rhythmischen Periodengrenzen zusammen; die Periodenstruktur ist typischerweise auf allen ihren vertikalen Niveaus rhythmisch invariant. Innerhalb der Takte bestehen Akzentvariationen, die aber ebenfalls eine hierarchische Untergliederung des Taktniveaus aufweisen. Da aber innerhalb eines Taktes die den Rhythmus definierende Akzentverteilung eine Invarianz höherer Ordnung als oberhalb des Taktniveaus darstellt, ist es verständlich, daß das kognitive System bei der Entnahme von Informationen unterhalb des Taktniveaus größeren Schwierigkeiten gegenübersteht als oberhalb des Taktniveaus: Wegen der höheren Invarianzordnung muß bei der Fourieranalyse ein zusätzlicher Dekompositionsschritt der Superpositionsfunktion stattfinden. Die Ursache für die schlechtere Differenzierung der Klicklokalisationshäufigkeiten und der Detektionszeiten für Klicks auf den unteren Niveaus könnte entweder darauf zurückzuführen sein, daß dieser Dekompositionsschritt nicht durchgeführt wurde, oder aber wegen der stärkeren Merkmalsvariation auf den unteren Niveaus fehleranfälliger ist und die Lokalisation hier eher einem Ratevorgang entspricht, was zu der beobachteten Gleichverteilung der Lokalisationshäufigkeiten auf Ton- und Fußniveau führen sollte.

Die Befunde von Experiment 3 können zwar ebenso wie die von Experiment 4 auf die aktive Antizipation höherer syntaktisch-rhyth-

misch definierter Grenzen innerhalb der musikalischen Periode zurückgeführt werden, aber diese Interpretation schließt keineswegs aus, daß die der Antizipation zugrundeliegende Repräsentationsstruktur hierarchisch organisiert ist. Ganz im Gegenteil: Eine die syntaktisch-rhythmisch markierten Grenzen antizipierende Form der kognitiven Verarbeitung ist nur deshalb möglich, weil die Repräsentationsstruktur bereits hierarchisch ist. Die Schlußfolgerung aus den Klicklokalisationsdaten, die bei der Wahrnehmung entstehende Repräsentation sei hierarchisch strukturiert, ist also durchaus gerechtfertigt.

Hätten wir in den Experimenten von Teil II Melodien ohne systematische Akzentverteilungen verwendet, dann wäre damit keineswegs sichergestellt gewesen, daß wir den Effekt des syntaktischen Wissens auf die Verarbeitung von Melodien in "reiner Form" hätten erfassen können. Es wäre dann nämlich wahrscheinlich gewesen, daß eine subjektive Rhythmisierung eingetreten wäre (Fraisse, 1978; Woodrow, 1951), die sowohl syntaktisch kompetente wie musikalisch unerfahrene Hörer vornehmen können. Diese subjektiven Rhythmen wären aber ebenso wie die im physikalischen Signal enthaltenden Akzentverteilungen perfekt mit der musikalischen Periodenstruktur korreliert, wenn die Hörer einen 2/2-Takt oder einen 4/4-Takt zugrunde legen würden, und damit für die aktive Antizipation struktureller Grenzen funktional den akustischen Akzenten vergleichbar. Hätten wir die für Musik typische Akzentvariation eleminiert, hätten wir einen wesentlichen Anteil der musikalischen Merkmalsvariationen unberücksichtigt gelassen und damit in erster Linie die Repräsentationsstruktur in einer für Musik atypischen Weise beeinflußt.

Versuch einer kognitionspsychologischen Beschreibung des strukturellen Hörens

Unser Ziel war es, strukturelles Hören kognitionspsychologisch zu beschreiben. Wieweit wir uns diesem Ziel haben nähern können, wollen wir nun untersuchen.

Auf der Grundlage der begrifflichen Analyse von Keil (1974a) konnten wir diese musikalische Höreinstellung folgendermaßen charakterisieren: Beim strukturellen Hören richtet der Hörer seine Aufmerksamkeit auf strukturelle Aspekte der Musik, deren zeitliche Erstreckung vom einzelnen Ton und dessen Merkmalen bis hin zum ganzen Musikstück reichen kann; dabei ist die Aufmerksamkeit nicht alleine auf die Identifikation einzelner struktureller Aspekte gerichtet, sondern ganz besonders auf das Erkennen der die einzelnen strukturellen Aspekte integrierenden Beziehungen im Gesamtzusammenhang eines Musikstücks, einer Klasse von Musikstücken des gleichen formalen Typs oder ähnlichen Stils.

Psychologisch von besonderem Interesse ist hierbei die Intention des Hörers, die Details im Zusammenhang übergeordneter Strukturen zu identifizieren, d.h. er muß seine Aufmerksamkeit auf strukturelle Details lenken können, ohne gleichzeitig die übergeordneten Zusammenhänge aus den Augen zu verlieren.

Ein lineares Modell, wie z.B. das von Massaro (1975), muß davon ausgehen, daß der Hörer vom einzelnen Ton oder von kleinen Tongruppen angefangen in einem konstruktiven Prozeß auf dem Niveau des Kurzzeitspeichers eine Repräsentation übergeordneter Zusammenhänge allmählich aufbaut. Dieser Kodierungsprozeß kann auch zu

einer hierarchischen Gedächtnisrepräsentation führen (Jones, 1978; Restle, 197o), aber die jeweils höheren Ebenen können erst dann kodiert werden, wenn a l l e niedrigeren Ebenen bereits vollständig kodiert sind. Entginge dem diesen Konstruktionsprozeß durchführenden Hörer auch nur ein einziger Ton, könnte er bereits sämtliche übergeordneten Beziehungsmuster nicht mehr richtig repräsentieren.

Wir haben stattdessen in Anlehnung an die Theorie der direkten Wahrnehmung von Gibson (1966a) und an die durch das Hierarchiemodell von Wolff (1977) gegebene Rahmenkonzeption ein strukturelles und funktionales Modell der auditiven Wahrnehmung von Tonfolgen und Musik entworfen, das von der Annahme einer primären Reizmusterrepräsentation ausgeht, die hierarchisch strukturiert ist. Diese hierarchische Struktur ist aufgrund der Systematik der Variation physischer Merkmale des Reizmusters bereits gegeben, so daß auf konstruktive Prozesse des Aufbaus einer Repräsentationsstruktur vollständig verzichtet werden kann. Der Entnahme von Informationen aus der Reizmusterrepräsentation sind strukturelle Grenzen gesetzt, die die Verarbeitungsprozesse, die auf Reizmusterrepräsentanzen zugreifen, reflektieren (vgl. auch Bregman, 1978a).

Funktionale Voraussetzung für einen Zugriff auf diese Repräsentationsstruktur ist ein die simultan als Superposition repräsentierte Verteilung physischer Merkmale in der Zeit schrittweise zergliedernder und intentional auf ein bestimmtes Repräsentationsniveau einstellbarer Fourieranalyseprozeß, der nach Prinzipien der Invarianzanalyse für eine gerade fokussierte Repräsentanz eine Binnengliederung erstellt, so daß die Merkmalsvariation zwischen den Binnengliedern niedriger als innerhalb ist. Während auf einem

bestimmten Repräsentationsniveau, das Binnenglieder gleicher Invarianzordnung und annähernd gleicher zeitlicher Erstreckung enthält, eine Repräsentanz lokal analysiert wird, bleibt das unmittelbar übergeordnete Niveau ebenso global repräsentiert, wie das unmittelbar untergeordnete, dessen Binnengliederung aufgrund der fortlaufend stattfindenden Fourieranalyse als solche repräsentiert ist, nicht aber die Identität einzelner Binnenglieder; Identitätsinformation ist erst bei Wechsel auf die untergeordnete Repräsentationsebene durch Fokussierung und Analyse eines Binnengliedes zu entnehmen.

Das Hören und Identifizieren zeitlich weitgespannter Beziehungen innerhalb eines Musikstücks ist durch Fokussierung entsprechend hoher Repräsentationsniveaus und durch Analyse seiner Binnenglieder möglich. Da bei Fokussierung eines Binnenglieds das übergeordnete Repräsentationsniveau weiterhin global repräsentiert bleibt, kann während der Binnengliedanalyse der zurückliegende Kontext berücksichtigt und der nachfolgende Kontext zumindest global antizipiert werden, d.h. in die strukturelle Detailanalyse fließen immer Informationen über übergeordnete Zusammenhänge ein, weshalb wir hier von einer heterarchischen Verarbeitung sprechen können (Turvey, 1977). Z.B. fließen bei der Analyse von Details des Moll-Themas in Ravels "Bolero" globale strukturelle Informationen über das vorausgehende Dur-Thema ein, so daß sich die melodischen Ähnlichkeiten beider Themen unmittelbar auffassen lassen.

Werden in schneller Folge mehrere Repräsentationsniveauwechsel in Richtung auf ein detailliertere Informationen enthaltendes Niveau

vorgenommen, dann ist bei jedem einzelnen Niveauwechsel zumindest der jeweils übergeordnete Kontext repräsentiert, so daß das schließlich fokussierte Detail einschließlich seiner vielschichtigen Einbettungen innerhalb der musikalischen Struktur identifiziert werden kann.

Da der Hörer nicht gleichzeitig mehrere Repräsentationsniveaus fokussieren kann, macht ein strukturelles Hören, das dem Hörer einen möglichst differenzierten Eindruck von der musikalischen Struktur vermitteln soll, eine sehr flexible Strategie des Repräsentationsniveauwechsels notwendig. Diese Flexibilität ist nur dann zu erreichen, wenn der Hörer die Niveauwechsel vorausplanen kann, indem er strukturell wichtige Stellen, nämlich solche, an denen auf mehreren Repräsentationsniveaus gleichzeitig Binnengliedgrenzen zusammenfallen (höhere Konstituentengrenzen), antizipiert, um zum richtigen Zeitpunkt das seinen Intentionen entsprechende Niveau fokussieren zu können.

Eine solche aktive Planung und Antizipation macht kognitive Strukturen notwendig, die aufgrund der aufgenommenen Informationen über die musikalische Struktur Extrapolationen in die Zukunft vornehmen, Identitätskategorien bereitstellen und den Fokussierungsprozeß so steuern, daß den Intentionen des Hörers gemäß die gewünschte Information einem Repräsentationsniveau entnehmbar wird. Solche kognitiven Schemata im Sinne Neisser's (1976) enthalten Spezifikationen über die in Musik zu erwartenden Strukturen, die als implizites, musikalisch-syntaktisches Wissen bezeichnet werden können. Ein differenziertes strukturelles Hören setzt also ein hohes Maß an Hörerfahrungen und ein hierdurch erlangtes differenziertes musikalisch-syntaktisches Wissen voraus.

Kognitive Prozesse bei nicht primär auf die musikalische Struktur gerichteten Höreinstellungen

Jede Form des Musikhörens ist auf kognitive Prozesse angewiesen, auch solche Höreinstellungen, die nicht primär auf das Erfassen musikalischer Strukturen gerichtet sind.

Musikalische Rezeptionsweisen unterscheiden sich 1. hinsichtlich der Differenziertheit, mit der die Reizmusterrepräsentation unter Verwendung kognitiver Schemata analysiert wird; die Differenziertheit ist sowohl von der Zahl als auch vom Niveau der berücksichtigten Repräsentationsebenen sowie von der Assimilations- und Akkommodationsfähigkeit der zur Verfügung stehenden kognitiven Schemata abhängig; sie unterscheiden sich 2. darin, daß bei einigen Höreinstellungen das Ergebnis der auf die Reizmusterrepräsentation zugreifenden Verarbeitungsprozesse bewußt repräsentiert ist, wie beim strukturellen Hören, während bei anderen Höreinstellungen die Gliederung der Reizmusterrepräsentation in dem Sinne zeitweise zumindest präattentiv bleibt, als nicht ständig eine intentionale Fokussierung eines Repräsentationsniveaus vorgenommen wird, wie dies beim "Hören" von Musik als klangliche Hintergrundkulisse der Fall ist; und 3. unterscheiden sie sich in der Art und Weise, in der die Reizmusterrepräsentation kategorial repräsentiert wird, d.h. sie unterscheiden sich darin, ob Musik als musikalische Struktur repräsentiert wird oder als "Zeichen" oder "Ausdruck" einer durch die klangliche Struktur kommunizierten Intention des Komponisten oder Interpreten, durch die die musikalische Struktur auf etwas außerhalb ihrer selbst liegendes verweist (semiotischer Aspekt), oder als "Auslöser" bzw. "Projektionsgrundlage" für die Empfindungen, Emotionen oder Erinnerungen des Hörers.

Kritisch für eine psychologische Differenzierung zwischen musikalischen Rezeptionsweisen ist also insbesondere die Art und Weise der Aufmerksamkeitssteuerung bei Zugriff auf die Repräsentationsstruktur sowie Inhalt und Differenziertheit der einem Hörer zur Verfügung stehenden kognitiven Schemata. Berücksichtigt werden muß aber außerdem noch der übergeordnete Handlungskontext, in dem der Musik eine die Handlung steuernde Funktion zukommt. Das Beispiel des Tanzes soll diesen Punkt verdeutlichen.

Beim Tanzen z.B. auf einer Party kann, wenn die verbale Kommunikation im Vordergrund steht, der Tanzmusik ein Minimum an Aufmerksamkeit zugewendet werden. Der Tänzer braucht lediglich innerhalb der ersten Takte eines Musikstücks den Grundrhythmus zu identifizieren, fokussiert also kurzzeitig das Taktniveau oder das ihm unmittelbar übergeordnete Niveau, spezifiziert aufgrund dieser Informationen ein kognitives Schema, das nun eine Extrapolation der rhythmischen Struktur in Form einer Antizipation vornimmt. Entweder bleibt ein den Rhythmus global repräsentierendes Niveau konstant fokussiert, oder aber dieses Niveau wird nur noch von Zeit zu Zeit fokussiert, um die Koordination der Bewegungen mit dem Rhythmus der Musik zu überprüfen.

Sobald aber eine Rhythmusänderung eintritt oder das Bewegungsprogramm intentional modifiziert werden soll, ist eine erneute Fokussierung eines Repräsentationsniveaus notwendig, das den Rhythmus identifizierbar werden läßt. Eine Ballettänzerin, deren Bewegungsabläufe nicht die monotone Periodik eines Partytanzes zeigen, muß dagegen ständig ihren Bewegungsablauf mit den häufiger wechselnden Rhythmen der Ballettmusik koordinieren. Sie muß ihren Bewegungsab-

lauf zeitlich weitreichend vorausplanen und deshalb rhythmische Strukturen sowohl auf niedrigen als auch auf hohen Repräsentationsniveaus fokussieren und dabei die Repräsentationshierarchie in der von Martin (1972) beschriebenen Weise von oben nach unten durchlaufen und an syntaktisch ausgezeichneten Stellen, die gleichfalls für den Bewegungsablauf den Übergang auf eine höhere Planungsebene verlangen, auf übergeordnete Niveaus zurückkehren.

Von vergleichbarer Komplexität ist die Handlungssteuerung eines Musikers. Auch hier müssen motorische Planungsebenen des als hierarchisch strukturiert aufzufassenden Handlungsprogramms (Turvey, 1977) mit allen strukturellen Niveaus des visuell repräsentierten Notentextes oder der Gedächtnisrepräsentationen koordiniert werden.

Auch solche als "emphatische Rezeption" oder "assoziativ-emotionale Rezeption" (Rauhe, Reinecke & Ribke, 1975) bezeichnete Höreinstellungen erfordern ein intentionales Fokussieren der Reizmusterrepräsentation. Da hierbei überwiegend strukturelle Merkmale sehr hoher Ordnung entnommen werden müssen, um Stimmungen oder den Ausdruck eines Musikstücks erfassen zu können, werden relativ hohe Repräsentationsniveaus fokussiert und nur selten Repräsentationsniveauwechsel vorgenommen. Die bei solchen Rezeptionsweisen die Fokussierung und Informationsentnahme steuernden kognitiven Schemata können als Spezifikationen von Merkmalskomplexen betrachtet werden, die für ein Individuum invariant mit einem bestimmten musikalischen Ausdruck verknüpft sind. Wie solche individuellen Invarianzkategorien entstehen, ist noch völlig ungeklärt. Ihre Aktivierung während eines Hörvorgangs ließe sich viel-

leicht analog zum Gedächtnismodell von Bransford et al. (1977) als Resonanzmechanismus begreifen.

Abschließende Bemerkung

Musikpsychologische Untersuchungen waren bisher mangels objektiver Methoden der strukturellen Beschreibung von Musik in besonderem Maße der Gefahr ausgesetzt, ihren Gegenstand aus den Augen zu verlieren, wenn das "Reizmaterial" den Notwendigkeiten experimenteller Rigorosität folgend so manipuliert wurde, daß die Realitätsnähe dieses Reizmaterials zur Musik unbestimmbar wurde. Es war unsere Absicht, im Teil II dieser Untersuchung "echtes" musikalisches Material zu verwenden, doch auch wir konnten uns nicht ganz der Notwendigkeit entziehen, Vereinfachungen der als Reizmaterial dienenden Melodien vorzunehmen, die wahrscheinlich dazu geführt haben, daß nicht alle Hörer diese Melodien als "echt" bezeichnen werden. Doch im Unterschied zu bisherigen Untersuchungen mit künstlich und willkürlich konstruierten Tonfolgen können wir aufgrund der Syntax exakt angeben, bezüglich welcher struktureller Aspekte sich die von uns verwandten Melodien von echten Volks- und Kinderliedern unterscheiden. Die Übertragbarkeit der berichteten Befunde auf das Hören echter Musik ist hierdurch gewiß eingeschränkt, aber die Aspekte der kognitiven Verarbeitung, auf die sich diese Einschränkungen beziehen, sind angebbar: Die Konstanthaltung rhythmischer Strukturen durch gleichlange Notenwerte und gleichlange Pausen zwischen Noten be-

wirkt eine besonders regelmäßige Reizmusterrepräsentation auf allen strukturellen Niveaus der Repräsentationshierarchie; die hierdurch entstehende Einschränkung der Merkmalsvariabilität innerhalb und zwischen Binnengliedern eines Niveaus bewirkt eine Konstanz der Fokussierungsschwierigkeit aufgrund einer konstant gehaltenen Deutlichkeit der Binnengliedgrenzen auf allen Repräsentationsniveaus.

ZUSAMMENFASSUNG

Ziel der vorliegenden Untersuchung ist die Formulierung und experimentelle Überprüfung eines strukturellen und funktionalen Modells der kognitiven Verarbeitung von Tonfolgen und Melodien. In Teil I wird die theoretische Rahmenkonzeption entwickelt und in Experimenten zur Wahrnehmung kurzer Tonfolgen überprüft, im Teil II wird diese Rahmenkonzeption dann sukzessive für die Wahrnehmung von Melodien spezifiziert und experimentell überprüft.

Ausgangspunkt für Teil I ist die Frage, ob es möglich ist, Aussagen über kognitive Verarbeitungsprozesse beim Hören von Musik zu machen, ohne gleichzeitig zu reflektieren, welche tonalen Strukturen spezifisch musikalische sind. Wir charakterisieren die Voraussetzungen des linearen Informationsverarbeitungsansatzes als solche, die eine Ausklammerung dieser Frage zu erlauben scheinen und prüfen die Tragfähigkeit dieser Voraussetzungen an den in der Literatur zur auditiven Mustererkennung berichteten Befunden. Ein großer Teil dieser Befunde läßt die Grundannahme linearer Informationsverarbeitungsmodelle, Wahrnehmung bestehe in der sukzessiven Analyse und Konstruktion von Repräsentationen aus elementaren Einheiten (Merkmalen, features), unwahrscheinlich erscheinen, weil auf dieser Grundlage nur wenige theoretisch befriedigende Erklärungen zur kognitiven Verarbeitung und Repräsentation von Tonfolgen zu geben sind. Zahlreiche Ergebnisse sprechen dafür, daß Merkmale höherer Ordnung, Intervallkonfigurationen oder Konturverläufe, zeitlich vor den sie konstituierenden tonalen Elementen repräsentiert sind.

Hieraus ergeben sich die für Teil I zentralen Fragestellungen: 1. Wie läßt sich die Identifikation der Struktur von Tonfolgen sowohl funktional als auch strukturell erklären, und 2. durch welches Modell der Repräsentation ist die phänomenale Ganzheitlichkeit der Wahrnehmung von Tonfolgen zu erklären?

Bei einem Überblick über neuere Befunde und Modelle zur visuellen Wahrnehmung komplexer Reizmuster finden wir unser Urteil bezüglich der Inadäquatheit des linearen Informationsverarbeitungsansatzes bestätigt, aber auch gleichzeitig theoretische Anknüpfungspunkte, die uns zur Formulierung eines Alternativmodells veranlassen, das zunächst in einer Spezifizierung des hierarchischen Repräsentationsmodells von Wolff (1977) für die Wahrnehmung von Tonfolgen besteht.

Das Hierarchiemodell geht davon aus, daß bereits die primäre Reizrepräsentation im echoischen Speicher eine hierarchische Struktur besitzt, die vertikal nach Niveaus unterschiedlicher Korngröße bzw. zeitlicher Erstreckung und horizontal in Verarbeitungseinheiten gegliedert ist, die bei Aufmerksamkeitszuwendung sequentiell fokussiert werden können. Es wird angenommen, daß Tonfolgen, die die psychische Präsenzzeit nicht überschreiten, in der Reihenfolge vom höchsten zum niedrigsten Repräsentationsniveau verarbeitet werden.

Wenn sich diese Annahme als begründet erweist, kann hierdurch erklärt werden, warum Tonfolgen zuerst phänomenal ganzheitlich repräsentiert sind. Durch die Annahme, daß im Verlauf eines die aufgabenrelevante Information entnehmenden Analyseprozesses ein

hierarchischer Einschachtelungsprozeß stattfindet, der die Repräsentationsebenen von oben nach unten durchläuft und die aufgabenrelevante Information dem jeweils höchsten Niveau entnimmt, das sie enthält, ist die strukturelle Identifikation von Tonfolgen in Abhängigkeit von der Konfiguration ihrer Intervalle zu erklären.

Dabei gehen wir zunächst davon aus, daß die Identitätskategorie eine "tonale Linie" (kontinuierlich auf- oder absteigende Tonfolge) ist, so daß im Verlauf des hierarchischen Einschachtelungsprozesses diejenige Repräsentationsebene aufgesucht werden muß, auf der alle Verarbeitungseinheiten solche "tonalen Linien" darstellen. Besteht die Gesamttonfolge bereits aus einer kontinuierlich auf- oder absteigenden Tonfolge, dann ist die Identifikation dieser Intervallkonfiguration bereits auf dem höchsten Repräsentationsniveau möglich, zeigt die Tonfolge aber Intervallrichtungswechsel, dann muß solange auf niedrigere Repräsentationsniveaus gewechselt werden, bis die einzelnen Verarbeitungseinheiten als "tonale Linien" identifizierbar sind.

Nach der Chunkingkonzeption von Wolff (1977), die im Gegensatz zu der Miller's (1956a) den funktionalen Ort des Chunkings auf das Niveau des sensorischen Speichers vorverlagert, sollte durch die Ähnlichkeit der Tonhöhen und/oder der zeitlichen Intervalle zwischen Tönen ein Chunkniveau gebildet werden, das die Entnahme der aufgabenrelevanten Information dann beschleunigt, wenn sie einem Chunk zu entnehmen ist, aber verlangsamt, wenn sie auf dem dem Chunkniveau untergeordneten Niveau entnommen werden muß.

Diese Hypothesen werden in Experiment 1 überprüft, indem die Reaktionszeit für die Identifikation der Richtung des letzten Intervalls einer Dreitonfolge gemessen wird. Variiert werden alle SOAs zwischen Tönen und die Intervallgrößen zwischen Tönen sowie die durch alle Töne gebildete Intervallkonfiguration (gleichsinnige Intervallkonfiguration bei konstanten Intervallrichtungen, gegensinnige Intervallkonfiguration bei Intervallrichtungswechsel). In den Kontrollbedingungen wird das kritische Intervall ohne den vorausgehenden ersten Ton dargeboten.

Der Effekt der Intervallkonfiguration stellt sich wie erwartet ein, aber die SOA- und Intervallgrößenvariationen zeigen nicht die aufgrund des Chunkingmodells zu erwartenden Effekte. Vielmehr wird durch die Daten deutlich, daß bei minimaler Variabilität der SOAs und Intervallgrößen innerhalb der Dreitonfolge die Identifikation am schnellsten verläuft. Obwohl der erste Ton nicht aufgabenrelevant ist und eigentlich unbeachtet bleiben könnte, übt er einen systematischen Einfluß auf die Identifikationsgeschwindigkeit durch seine Einbettung in die Konfiguration aller drei Töne aus.

Wir ziehen aus diesem Befundmuster die Schlußfolgerung, daß die Repräsentation nicht in diskreten Verarbeitungseinheiten, sondern in einer der Merkmalsvariction über die gesamte Tonfolge analogen Form besteht und geben die Chunkingkonzeption auf. Stattdessen nehmen wir an, daß Invarianzen (Gibson, 1966a), in die die Gesamtinformation über die Variabilität aller Merkmale simultan eingehen, Binnengliedor auf den einzelnen Repräsentationsniveaus

ausgliedern. Eine Invarianzanalyse wird durch eine schrittweise Dekomposition der als Superposition aller Merkmalsverteilungen über die Zeit repräsentierten Tonfolge aufgrund eines Fourieranalyseprozesses (Pibram, 1977a,b) geleistet, der die Binnengliederung des jeweils zu fokussierenden Niveaus in Perioden gleicher Invarianzordnung erstellt.

Da in diesem Modell die Dekomposition der Superposition aller Merkmalsverteilungen um so mehr Schritte benötigt, je mehr Merkmale gleichzeitig innerhalb einer Tonfolge variieren, sollte die Identifikation einer Tonfolge um so länger dauern, je größer die Dimensionalität ihrer Merkmalsvariation ist, und da ein Repräsentationsebenenwechsel Zeit benötigt, sollte die Identifikation der Gesamttonfolge schneller sein als die eines Teils aus ihr, aber die Identifikationszeit für einen Teil sollte gleichzeitig von der Intervallkonfiguration abhängen, in die dieses Teil eingebettet ist.

Diese Hypothesen werden von den Daten aus Experiment 2 eindeutig unterstützt. Im Experiment werden Viertonfolgen dargeboten und mit einer visuellen Identifikationskategorie verglichen, die die Intervallkonfiguration spezifiziert (same/different-Vergleiche mit Reaktionszeitmessung). Bei Identifikation der Gesamttonfolge ist die Reaktionszeit kürzer als bei Identifikation der ersten oder zweiten Intervallrichtung. Die am schnellsten zu identifizierende Intervallkonfiguration ist, wie auch schon in Experiment 1, die kontinuierlich auf- oder absteigende Tonfolge, die zweitschnellste diejenige, in der von einem zum anderen Ton die Intervallrichtung wechselt (hier besteht aber Invarianz der Inter-

vallrichtungen zwischen 1. und 2. Tonpaar) und die langsamste ist diejenige, die nur einen Intervallrichtungswechsel enthält (keine Invarianz der Intervallrichtung zwischen Tonpaaren). Dieser Befund stimmt mit der Invarianzkonzeption überein, steht aber im Gegensatz zu der Vermutung von Ortmann (1933), wonach die Identifikationsschwierigkeit eine Funktion der Anzahl der Intervallrichtungswechsel sein sollte.

Die in Teil I entwickelte theoretische Rahmenkonzeption wird in Teil II zur Erklärung kognitiver Prozesse bei der Wahrnehmung von Melodien angewendet, die mit Hilfe einer Syntax deutscher Volks- und Kinderlieder generiert werden. Die in Teil II verfolgte Fragestellung ist die, ob eine Beziehung zwischen der Repräsentationsstruktur beim Hören von Melodien und ihrer syntaktischen Struktur nachweisbar ist.

In Experiment 3 wird diese Frage mit syntaktisch trainierten Vpn mit der aus der Psycholinguistik stammenden Methode der Klicklokalisation zu beantworten versucht. Die Lokalisationshäufigkeiten und die Fehllokalisationsbeträge und -richtungen zeigen im Prinzip die aus den psycholinguistischen Untersuchungen bekannten Verteilungen, die deutliche Hinweise auf eine hierarchische Repräsentationsstruktur enthalten, die mit der syntaktischen Struktur der musikalischen Perioden unserer Melodien übereinstimmt.

In Experiment 4 untersuchen wir dieselbe Frage mittels einer Klickdetektionsaufgabe, wobei zusätzlich ein Vergleich zwischen syntaktisch trainierten und untrainierten Hörern vorgenommen wird. Die Detektionszeiten zeigen wieder im Prinzip die von der syntaktischen Struktur abhängigen Effekte, allerdings nur bei den

syntaktisch trainierten Hörern in deutlicher Form. Die untrainierten Hörer reagieren dagegen in der zweiten Hälfte einer Melodie unabhängig von ihrer syntaktischen Struktur. Wir interpretieren diesen Unterschied zwischen den beiden Hörergruppen als Hinweis auf eine aktive Antizipation von syntaktisch-rhythmisch ausgezeichneten Stellen einer Melodie durch die trainierten Hörer, die ihnen aufgrund ihres syntaktischen Wissens ermöglicht wird, während die untrainierten Hörer schnell den Überblick über die musikalische Struktur verlieren und damit nicht mehr zu Antizipationen fähig sind.

Experiment 5 überprüft Hypothesen über die bei Zugriff auf eine hierarchische Repräsentationsstruktur ablaufenden Verarbeitungsprozesse. Es wird im wesentlichen die Hypothese geprüft, daß das Erkennen einer Beziehung zwischen zwei syntaktisch definierten Ausschnitten einer Melodie von der syntaktischen Nähe (die Ausschnitte gehören auf dem unmittelbar übergeordneten Strukturniveau zur selben oder zu verschiedenen syntaktischen Einheiten) und von der zeitlichen Nähe abhängig ist.

Die Reaktionszeiten für same/different-Vergleiche zwischen syntaktischen Einheiten auf drei verschiedenen Strukturniveaus bestätigen diese Erwartung: Syntaktisch nahe Ausschnitte können wesentlich schneller richtig miteinander verglichen werden als syntaktisch ferne, obwohl die zeitliche Nähe in beiden Fällen dieselbe ist. Dieser Befund ist u.E. der deutlichste Beleg dafür, daß die Repräsentation von Melodien, deren syntaktische Beschreibung eine hierarchische Struktur offenbart, hierarchisch strukturiert ist.

In der abschließenden Diskussion einiger Aspekte der Befunde und des Modells beurteilen wir die kognitive Adäquatheit der den Melodien zugrundeliegenden syntaktischen Beschreibung und diskutieren die funktionale Rolle musikalisch-syntaktischen Wissens für die kognitive Verarbeitung beim Hören von Musik. Abschließend versuchen wir, die als "strukturelles Hören" (Keil, 1974a) bezeichnete musikalische Rezeptionsweise aufgrund unseres Modells kognitionspsychologisch zu beschreiben und Unterschiede in der kognitiven Verarbeitung von Musik zwischen dieser und anderen Rezeptionsweisen herauszuarbeiten, in denen das Erkennen struktureller Zusammenhänge in der gehörten Musik nicht so stark im Vordergrund steht.

L I T E R A T U R V E R Z E I C H N I S

Abraham, L.K. & Dahlhaus, C., Melodielehre. Köln: Gerig, 1972

Abrams, K. & Bever, T.G., Syntactic structure modifies attention during speech perception and recognition. Quarterly Journal of Experimental Psychology, 1969, 21, 280 - 290

Adorno, T.W., Fragment über Musik und Sprache. In: Jahresring 1956/57. Stuttgart, 1956 (zit. nach Knaus, 1973)

Adorno, T.W., Einleitung in die Musiksoziologie. Reinbeck b. Hamburg: Rowohlt, 1968

Alt, M.,Die Erziehung zum Musikhören. Leipzig, 1935

Arom, S., Essai d'une notation des monodies à des fins d'analyse. Revue de Musicologie, 1970, 55, 172 - 216

Asch, S.E., A problem in the theory of associations. Psychologische Beiträge, 1962, 6, 553 - 563

Asch, S.E., Ceraso, J. & Heimer, W., Perceptual conditions of association. Psychological Monographs, 1960, 74, No. 3

Attneave, F., Informationstheorie in der Psychologie. Bern und Stuttgart: Huber, 1965

Attneave, F. & Olson, R.K., Pitch as a medium: A new approach to psychophysical scaling. American Journal of Psychology, 1971, 84, 147 - 165

Axelrod, S. & Guzy, L.T., Underestimation of dichotic click rates: Results using methods of absolute estimation and constant stimuli. Psychonomic Science, 1968, 12, 133 - 134

Baddeley, A.D., Immediate memory and the 'perception' of letter sequences. Quarterly Journal of Experimental Psychology, 1964, 16, 364 - 367

Baddeley, A.D., The trouble with levels: A reexamination of Craik and Lockhart's framework for memory research. Psychological Review, 1978, 85, 139 - 152

Bamber, D., Reaction times and error rates for "same"-"different" judgments of multidimensional stimuli. Perception & Psychophysics, 1969, 6, 169 - 174

Banks, W.P. & Prinzmetal, W., Configural effects in visual information processing. Perception & Psychophysics, 1976, 19, 361 - 367

Baroni, M. & Jacoboni, C., Proposal for a grammar of melody. The Bach chorals. Montréal: Les Presses de d'Université de Montréal, 1978

Bean, C., Information theory analyses of four sonata expositions. Journal of Music Theory, 1966, 10, 97 - 137

Bell, H.H. & Handel, S., The role of pattern goodness in the reproduction of backward masked patterns. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 139 - 150

Bernstein, L., Musik - Die offene Frage. München: Molden, 1979

Bernstein, N., The coordination and regulation of movements. London: Pergamon Press, 1967

Bertelson, P. & Tisseyre, F., Perceiving the sequence of speech and non-speech stimuli. Quarterly Journal of Experimental Psychology, 1970, 22, 653 - 662

Bever, T.G., Serial position and response biases do not account for the effect of syntactic structure on the location of brief noises during sentences. Journal of Psycholinguistic Research, 1973, 2, 287 - 288

Bever, T.G. & Chiarello, R.J., Cerebral dominance in musicans and nonmusicans. Science, 1974, 185, 537 - 539

Bever, T.G., Lackner, J.R. & Kirk, R., The underlying structures of sentences are the primary units of immediate speech processing. Perception & Psychophysics, 1969, 5, 225 - 234

Bever, T.G., Lackner, J.R. & Stolz, W., Transitional probability is not a general mechanism for the segmentation of speech. Journal of Experimental Psychology, 1969, 79, 387 - 394

Biederman, I., Perceiving real-world scenes. Science, 1972, 177, 77 - 79

Biederman, I., Glass, A.L. & Stacey, E.W., Searching for objects in real-world scenes. Journal of Experimental Psychology, 1973, 97, 22 - 27

Biederman, I., Rabinowitz, J.C., Glass, A.L. & Stacey, E.W., On the information extracted from a glance at a scene. Journal of Experimental Psychology, 1974, 103, 597 - 600

Bimberg, S., Zur Festlegung des Funktionsbegriffs in der Musik. Wissenschaftliche Zeitschrift der Martin-Luther-Universität Halle-Wittenberg, 1954, 3, 737 - 740

Bobrow, D.G. & Norman, D.A., Some principles of memory schemata. In: D.G. Bobrow & A. Collins (Eds.), Representation and understanding. New York: Academic Press, 1975

Boilés, C., Tepehua thought-song: A case of semantic signalling. Ethnomusicology, 1967, 11, 267 - 292

Bolton, T.L., Rhythm. American Journal of Psychology, 1894, 6, 145 - 238

Bower, G.H., Organizational factors in memory. Cognitive Psychology, 1970, 1, 18 - 46

Bower, G.H., Stimulus-sampling theory of encoding variability. In: A.W. Melton & E. Martin (Eds.), Coding processes in human memory. Washington, D.C.: Winston, 1972

Bower, G.H. & Glass, A.L., Structural units and the redintegrative power of picture fragments. Journal of Experimental Psychology: Human Learning and Memory, 1976, 32, 109 - 118

Bower, G.H. & Springston, F., Pauses as recoding points in letter series. Journal of Experimental Psychology, 1970, 83, 421 - 430

Bower, G.H. & Winzenz, D., Group structure, coding, and memory for digit series. Journal of Experimental Psychology Monographs, 1969, 80, No.2, Part 2, 1 - 17

Box, G.E.P., Non-normality and tests of variance. Biometrica, 1953, 40, 318 - 335

Bransford, J.D., McCarrell, N.S., Franks, J.J. & Nitsch, K.E., Toward unexplaining memory. In: R. Shaw & J.D. Bransford (Eds.), Perceiving, acting, and knowing. Toward an ecological psychology. Hillsdale, N.J.: Erlbaum, 1977

Bregman, A.S., The formation of auditory streams. In: J. Requin (Ed.), Attention and performance VII. Hillsdale, N.J.: Erlbaum, 1978 (a)

Bregman, A.S., Auditory streaming: Competition among alternative organizations. Perception & Psychophysics, 1978, 23, 391 - 398 (b)

Bregman, A.S. & Campbell, J., Primary auditory stream segregation and perception of order in rapid sequences of tones. Journal of Experimental Psychology, 1971, 89, 244 - 249

Bregman, A.S. & Dannenbring, G.L., The effect of continuity on auditory stream segregation. Perception & Psychophysics, 1973, 13, 308 - 312

Bregman, A.S. & Rudnicky, A.I., Auditory segregation: Stream or streams? Journal of Experimental Psychology: Human Perception and Performance, 1975, 1, 263 - 267

Breitmeyer, B.G., Simple reaction time as a measure of the temporal response properties of transient and sustained channels. Vision Research, 1975, 15, 1411 - 1412

Bright, W., Language and music: Areas for cooperation. Ethnomusicology, 1963, 7, 26 - 32

Brincker, J., Statistical analysis of music. An application of information theory. Svensk Tidskrift för Musikforsking, 1970, 52, 53 - 57

Broadbent, D.E., Perception and communication. London/New York: Pergamon Press, 1958

Broadbent, D.E., Stimulus set and response set: Two kinds of selective attention. In: D.J. Mostofsky (Ed.), Attention: Contemporary theory and analysis. New York: Appleton-Century-Crofts, 1970

Broadbent, D.E., Waves in the eye and ear. Journal of Sound and Vibration, 1975, 41, 113 - 125

Broadbent, D.E., The hidden preattentive processes. American Psychologist, 1977, 32, 109 - 118

Broadbent, D.E. & Ladefoged, P.N., Letter to the editor: Auditory perception of temporal order. Journal of the Acoustical Society of America, 1959, 31, 1539

Bruner, J.S., On perceptual readiness. Psychological Review, 1957, 64, 123 - 152

Bünting, K.-D., Einführung in die Linguistik. Frankfurt: Athenäum, 1972

Bukofzer, M., Musique in the Baroque era. New York: Norton, 1947

Burns, E.M. & Ward, W.D., Categorial perception of musical intervals. Journal of the Acoustical Society of America, 1973, 54, S 96

Campbell, D.T. & Stanley, J.C., Experimental and quasi-experimental designs for research. Chicago: Rand, 1963

Campbell, F.W., The transmission of spatial information through the visual system. In: F.O. Schmitt & F.G. Worden (Eds.), The neurosciences third study program. Cambridge, Mass.: MIT-Press, 1974

Capurso, A., Fisichelli, V.R., Gilman, L., Gutheil, E.A., Wright, J.T. & Paperte, F., Music and your emotions. New York, 1972

Carr, T.H. & Bacharach, V.R., Perceptual tuning and conscious attention: Systems of input regulation in visual information processing. Cognition, 1976, 4, 281 - 302

Carroll, J.M. & Bever, T.G., Segmentation in cinema perception. Science, 1975, 191, 1053 - 1055

Carterette, E.C. & Friedman, M.P. (Eds.), Handbook of perception, vol. IV: Hearing. New York: Academic Press, 1978

Chapin, P.G., Smith, T.S. & Abrahamson, A.A., Two factors in perceptual segmentation of speech. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 164 - 173

Cherry, E.C., Some experiments on the recognition of speech, with one and with two ears. Journal of the Acoustical Society of America, 1953, 25, 975 - 979

Cherry, E.C. & Taylor, W.K., Some further experiments on the recognition of speech with one and two ears. Journal of the Acoustical Society of America, 1954, 26, 554 - 559

Chomsky, N., Three models for the description of language. IRE Transactions on Information Theory, 1956, IT-2, 113 - 124

Chomsky, N., Syntactic structures. Den Haag: Mouton, 1957

Chomsky, N., On the notion "rule of grammar". In: J.A. Fodor & J.J. Katz (Eds.), The structure of language: Readings in the philosophy of language. Englewood Cliffs, N.J.: Prentice-Hall, 1964

Chomsky, N., Aspects of a theory of syntax. Cambridge, Mass.: MIT-Press, 1965

Chomsky, N., Language and mind. New York: Hartcourt, Brace and World, 1968

Chomsky, N. & Halle, M., The sound pattern of English. New York: Harper & Row, 1968

Cohen, J.E., Information theory and music. Behavioral Science, 1962, 7, 137 - 163

Conrad, R., Very brief delays of immediate recall. Quarterly Journal of Experimental Psychology, 1960, 12, 45 - 47

Coons, E. & Kraehenbuehl, D., Information as a measure of structure in music. Journal of Music Theory, 1957, 1, 127 - 161

Cooper, G.W. & Meyer, L.B., The rhythmic structure of music. Chicago: University of Chicago Press, 1960

Cooper, L.A., Recent themes in visual information processing: A selected overview. In: R.S. Nickerson (Ed.), Attention and performance VIII. Hillsdale, N.J.: Erlbaum, 1980

Cooper, R., Propositions pour un modèle transformationnel de description musicale. Musique en jeu, 1973, 10, 70 - 88

Cooper, R., Abstract structure and the Indian Rāga system. Ethnomusicology, 1977, 21, 1 - 32

Craig, J.D., Asymmetries in processing auditory nonverbal stimuli. Psychological Bulletin, 1979, 86, 1339 - 1349

Craik, F.I.M. & Lockhart, R.S., Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 671 - 684

Creel, W., Boomsliter, P.C. & Powers, S.R., Sensation of tone as perceptual forms. Psychological Review, 1970, 77, 534 - 545

Crickmore, L., The musical Gestalt. Music Review, 1972, 33, 285 - 291

Crowder, R.G., Representation of speech sounds in precategorial acoustic storage. Journal of Experimental Psychology, 1973, 98, 14 -24

Crowder, R.G., Mechanisms of auditory backward masking in the stimulus suffix effect. Psychological Review, 1978, 85, 502 - 524

Crowder, R.G. & Morton, J., Precategorial acoustic storage (PAS). Perception & Psychophysics, 1969, 5, 365 - 373

Crowder, R.G. & Raeburn, V.P., The stimulus suffix effect with reversed speech. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 342 - 345

Cuddy, L.L. & Cohen, A.J., Recognition of transposed melodic sequences. Quarterly Journal of Experimental Psychology, 1976, 28, 255 - 270

Dalrymple-Alford, E.C., Response bias and judgement of the location of clicks in sentences. Perception & Psychophysics, 1976, 19, 303 - 308

Dannenbring, G.L. & Bregman, A.S., Stream segregation and the illusion of overlap. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 544 - 555 (a)

Dannenbring, G.L. & Bregman, A.S., The effect of silence between tones on auditory stream segregation. Journal of the Acoustical Society of America, 1976, 59, 987 - 989 (b)

Darwin, C.J. & Baddeley, A.D., Acoustic memory and the perception of speech. Cognitive Psychology, 1974, 6, 41 - 60

Darwin, C.J., Turvey, M.T. & Crowder, R.G., An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. Cognitive Psychology, 1972, 3, 255 - 267

Davies, J.B., The psychology of music. London: Hutchinson, 1978

Davies, J.B., Memory for melodies and tonal sequences: A theoretical note. British Journal of Psychology, 1979, 70, 205 - 210

Davies, J.B. & Jennings, J., Reproduction of familiar melodies and the perception of tonal sequences. Journal of the Acoustical Society of America, 1977, 61, 534 - 541

Davies, J.B. & Yelland, A., Effects of two training procedures on the production of melodic contour, in short-term memory for tonal sequences. Psychology of Music, 1977, 5, 3 - 9

Deese, J., The structure of associations in language and thought. Baltimore: Hopkins, 1965

Deutsch, D., Music recognition. Psychological Review, 1969, 76, 300 - 307

Deutsch, D., Tones and numbers: Specificity of interference in short-term memory. Science, 1970, 168, 1604 - 1605

Deutsch, D., Octave generalization and tune recognition. Perception & Psychophysi 1972, 11, 411 - 412

Deutsch, D., Auditory memory. Canadian Journal of Psychology, 1975, 29, 87 - 105 (a)

Deutsch, D., The organization of short-term memory for a single acoustic attribute. In: D. Deutsch & J.A. Deutsch (Eds.), Short-term memory. New York: Academic Press, 1975 (b)

Deutsch, D. The psychology of music. In: E.C. Carterette & M.P. Friedman (Eds.), Handbook of perception, vol. X: Perceptual ecology. New York: Academic Press, 1978

Deutsch, D., The octave illusion and the what-where connection. In: R.S. Nickerson (Ed.), Attention and performance VIII. Hillsdale, N.J.: Erlbaum, 1980 (a)

Deutsch, D., The processing of structured and unstructured tonal sequences. Perception & Psychophysics, 1980, 28, 381 - 389 (b)

Deutsch, J.A. & Deutsch, D., Attention: Some theoretical considerations. Psychological Review, 1963, 70, 80 - 90

Dewar, K.M., Cuddy, L.L. & Mewhort, D.J.K., Recognition memory for single tones with and without context. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 60 - 67

Divenyi, P.L. & Hirsh, I.J., Identification of the order of three tones. Journal of the Acoustical Society of America, 1973, 54, 315 - 316

Divenyi, P.L. & Hirsh, I.J., Identification of temporal order in three-tone sequences. Journal of the Acoustical Society of America, 1974, 55, 144 - 151

Divenyi, P.L. & Hirsh, I.J., The effect of blanking on the identification of temporal order in three-tone sequences. Perception & Psychophysics, 1975, 17, 246 - 252

Dooling, D.J., Rhythm and syntax in sentence perception. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 255 - 264

Dorman, M.F., Cutting, J.E. & Raphael, L.J., Identification of vowel order: Concaterated versus formant connected sequences. Journal of Experimental Psychology: Human Perception and Performance, 1975, 1, 121 - 129

Doughty, J.M. & Garner, W.R., Pitch characteristics of short tones, two kinds of pitch threshold. Journal of Experimental Psychology, 1947, 37, 351 - 365

Dowling, W.J., Recognition of inversions of melodies and melodic contours. Perception & Psychophysics, 1971, 9, 348 - 349

Dowling, W.J., Recognition of melodic transformations: Inversion, retrograde, and retrograde inversion. Perception & Psychophysics, 1972, 12, 417 - 421

Dowling, W.J., Rhythmic groups and subjective chunks in memory for melodies. Perception & Psychophysics, 1973, 14, 37 - 40 (a)

Dowling, W.J., The perception of interleaved melodies. Cognitive Psychology, 1973, 5, 322 - 337 (b)

Dowling, W.J., Scale and contour: Two components of a theory of memory for melodies. Psychological Review, 1978, 85, 341 - 354

Dowling, W.J. & Fujitani, D.S., Contour, interval, and pitch recognition in memory for melodies. Journal of the Acoustical Society of America, 1971, 49, 524 - 531

Dowling, W.J. & Hollombe, A.W., The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour. Perception & Psychophysics, 1977, 21, 60 - 64

Edmonston, W.E., Familiarity and musical training in the esthetic evaluation of music. Journal of Social Psychology, 1969, 79, 109 - 111

Efron, R., The effect of handedness on the perception of simultaneity and temporal order. Brain, 1963, 86, 261 - 284

Efron, R., Conservation of temporal information by perceptual systems. Perception & Psychophysics, 1973, 14, 518 - 530

Egan, D.E. & Greeno, J.G., Theory of rule induction: Knowledge acquired in concept learning, serial pattern learning, and problem solving. In: L.W. Gregg (Ed.), Knowledge and cognition. Potomac, Md.: Erlbaum, 1974

Egan, J.P., Carterette, E.C. & Thwing, E.J., Some factors affecting multi-channel listening. Journal of the Acoustical Society of America, 1954, 26, 774 - 782

Ehrenfels, C.v., Über Gestaltqualitäten. Vierteljahresschrift zur Wissenschaftlichen Philosophie, 1890, 14

Engelkamp, J., Psycholinguistik. München: Fink, 1974

Estes, W.K., An associative basis for coding and organization in memory. In: A.W. Melton & E. Martin (Eds.), Coding processes in human memory. Washington, D.C.: Winston, 1972 (a)

Estes, W.K., Interactions of signal and background variables in visual processing. Perception & Psychophysics, 1972, 12, 278 - 286 (b)

Eysenck, M.W., Levels of processing: A critique. British Journal of Psychology, 1978, 69, 157 - 169

Falk, G.,Interpretation of imperfect line data as a three dimensional scene. Artificial Intelligence, 1972, 3, 101 - 144

Faltin, P., Der Verstehensbegriff im Bereich des Ästhetischen. In: P. Faltin & H.-P. Reinecke (Hg.), Musik und Verstehen. Köln: Volk, 1973

Farnsworth, P.R., The social psychology of music. New York, 1958

Fay, T., Perceived hierarchic structure in language and music. Journal of Music Theory, 1971, 15, 113 - 137

Fay, T., Context analysis of musical gestures. Journal of Music Theory, 1974, 18, 125 - 151

Flavell, J.H. & Draguns, J., A microgenetic approach to perception and thought. Psychological Bulletin, 1957, 54, 197 - 217

Fodor, J.A. & Bever, T.G., The psychological reality of linguistic segments. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 414 - 420

Fodor, J.A., Bever, T.G. & Garrett, M.F., The psychology of language. New York: McGraw-Hill, 1974

Foreit, K.G., Short-lived auditory memory for pitch. Perception & Psychophysics, 1976, 19, 368 - 370

Fraisse, P., Les structures rythmiques. Louvain: Studia Psychologica, 1956

Fraisse, P., Time and rhythm perception. In: E.C. Carterette & M.P. Friedman (Eds.), Handbook of perception, vol. VIII. New York: Academic Press, 1978

Francès, R., La perception de la musique. Paris, 1958

Fucks, W., Mathematische Musikanalyse und Randomfolgen. Musik und Zufall. Gravesaner Blätter, 1962, 6, 132 - 145

Fucks, W., Gibt es mathematische Gesetze in Sprache und Musik? In: H. Frank (Hg.), Kybernetik - Brücke zwischen den Wissenschaften. Darmstadt, 1965

Fucks, W., Nach allen Regeln der Kunst. Stuttgart, 1968

Fucks, W. & Lauter, J., Exaktwissenschaftliche Musikanalyse. Köln und Opladen: Westdeutscher Verlag, 1965

Gardiner, J.M., Levels of processing in word recognition and subsequent free recall. Journal of Experimental Psychology, 1974, 102, 101 - 105

Garner, W.R., Uncertainty and structure as psychological concepts. New York: Wiley, 1962

Garner, W.R., The processing of information and structure. Potomac, Md.: Erlbaum, 1974

Garner, W.R., Aspects of a stimulus: Features, dimensions, and configurations. In: E. Rosch & B.B. Lloyd (Eds.), Cognition and categorization. Hillsdale, N.J.: Erlbaum, 1978

Garner, W.R., Association lecture: Functional aspects of information processing. In: R.S. Nickerson (Ed.), Attention and performance VIII. Hillsdale, N.J.: Erlbaum, 1980

Garner, W.R. & Clement, D.E., Goodness of pattern and pattern uncertainty. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 446 - 452

Garner, W.R. & Gottwald, R.L., Some perceptual factors in the learning of sequential patterns of binary events. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 582 - 589

Garner, W.R. & Gottwald, R.L., The perception and learning of temporal patterns. Quarterly Journal of Experimental Psychology, 1968, 20, 97 - 109

Garrett, M., Bever, T.G. & Fodor, J.A., The active use of grammar in speech perception. Perception & Psychophysics, 1965, 1, 30 - 32

Gibson, E.J., Principles of perceptual learning and development. Englewood Cliffs, N.J.: Prentice-Hall, 1969

Gibson, J.J., The perception of the visual world. Boston: Houghton Mifflin, 1950

Gibson, J.J., The senses considered as perceptual systems. Boston: Houghton Mifflin, 1966 (a)

Gibson, J.J., The problem of temporal order in stimulation and perception. Journal of Psychology, 1966, 62, 141 - 149

Gibson, J.J., The theory of affordances. In: R. Shaw & J. Bransford (Eds.), Perceiving, acting, and knowing. Toward an ecological psychology. Hillsdale, N.J.: Erlbaum, 1977

Gibson, J.J. & Gibson, E.J., Perceptual learning: Differentiation or enrichment? Psychological Review, 1955, 62, 32 - 41

Ginsburg, A., The perception of visual form: A two-dimensional filter analysis. In: V.D. Gleger (Ed.), Proceedings of the fourth symposium on sensory system physiology. Leningrad: Russische Akademie der Wissenschaften, 1976 (zit. nach Kinchla & Wolfe, 1979)

Green, D.M., Temporal auditory acuity. Psychological Review, 1971, 78, 540 - 551

Green, D.M., Minimum integration time. In: A.R. Møller (Ed.), Basic mechanisms in hearing. New York: Academic Press, 1973

Greene, P.H., Introduction. In: I.M. Gelfand, V.S.Gurfinkel, S.V. Fomin & M.L. Tsetlin (Eds.), Models of the structural-functional organization of certain biological systems. Cambridge, Mass.: MIT-Press, 1971 (a) (zit. nach Turvey, 1977)

Greene, P.H., Problems of organization of motor systems. Quarterly Report No. 29, Institute for Computer Research. Chicago: University of Chicago, 1971 (b) (zit. nach Turvey, 1977)

Greenhouse, S. W. & Geisser, S., On methods in the analysis of profile data. Psychometrica, 1959, 24, 95 - 112

Gregg, F.M., Materializing the ghost of Köhler's Gestalt Psychology. Psychological Review, 1932, 39, 257 - 270

Gregory, A.H., Perception of clicks in music. Perception & Psychophysics, 1978, 24, 171 - 174

Gregory, R.L., Seeing as thinking. An active theory of perception. London Times Literary Supplement, June 23, 1972, 707 - 708

Guilford, J.P., Personality. New York: McGraw-Hill, 1959

Guzy, L.T. & Axelrod, S., Interaural attention shifting as response. Journal of Experimental Psychology, 1972, 95, 290 - 294

Haber, R.N., Perception and thought: An information processing analysis. In: J.F. Voss (Ed.), Approaches to thought. Pittsburg: University of Pittsburg Press, 1967

Hacker, W., Zur Anforderungsabhängigkeit der Nutzung hierarchischer Ordnung in Sequenzen. Zeitschrift für Psychologie, 1977, 185, 1 - 33

Halperin, Y., Nachshon, I. & Carmon, A., Shift of ear superiority in dichotic listening to temporal patterend nonverbal stimuli. Journal of the Acoustical Society of America, 1973, 53, 46 - 49

Handel, S., Temporal segmentation of repeating auditory patterns. Journal of Experimental Psychology, 1973, 101, 46 - 54

Handel, S., Perceiving melodic and rhythmic auditory patterns. Journal of Experimental Psychology, 1974, 103, 922 - 933

Handel, S. & Pickens, J.D., The effect of response constraints on the perception of repeating auditory patterns. Bulletin of the Psychonomic Society, 1977, 9, 253 - 255

Handel, S. & Yoder, D., The effects of intensity and interval rhythms on the perception of auditory and visual patterns. Quarterly Journal of Experimental Psychology, 1975, 27, 111 - 122

Harweg, R., Noch einmal: Sprache und Musik. Poetica, 1967, 1, 536 - 566

Harweg, R., Language and music: An immanent and sign theoretic approach. Foundations of Language, 1968, 4, 270 - 281

Harweg, R., Kann man Musik verstehen? International Review of the Aesthetics and Sociology of Music, 1972, 3, 173 - 186

Hebb, D.O., The organization of behavior. New York: Wiley, 1949

Heise, G.A. & Miller, G.A., An experimental study of auditory patterns. American Journal of Psychology, 1951, 64, 68 - 77

Helmholtz, H.v., Die Lehre von den Tonempfindungen als physiologische Grundlage für eine Theorie der Musik (1863), Nachdruck: Braunschweig, 1963

Henning, G.B., Hertz, B.G. & Broadbent, D.E., Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequencies. Vision Research, 1975, 15, 887 - 898

Herrmann, T., Der Schema-Begriff in der Denkpsychologie. Bericht vom Kongreß der Deutschen Gesellschaft für Psychologie, 1959, 22, 115 - 132

Herrmann, T., Die Psychologie und ihre Forschungsprogramme. Göttingen: Hogrefe, 1976

Hevner, K., Tests for the aesthetic appreciation in the field of music. Journal of Applied Psychology, 1930, 14, 470 - 477

Hevner, K., The affective value of certain elements of musical form. Psychological Bulletin, 1934, 31, 678 - 679

Hevner, K., Expression in music: A discussion of experimental studies and theories. Psychological Review, 1935, 47, 186 - 204

Hevner, K., Experimental studies of the elements of expression in music. American Journal of Psychology, 1936, 48, 246 - 268

Hevner, K., The affective value of pitch and tempo in music. American Journal of Psychology, 1937, 49, 621 - 630

Hick, W.E., On the rate of gain of information. Quarterly Journal of Experimental Psychology, 1952, 4, 11 - 26

Hiller, L.A., Informationstheorie und Musik. Darmstädter Beiträge zur Neuen Musik, Bd. 3. Mainz: Schott, 1964

Hiller, L.A. & Fuller, R., Structure and information in Webern's symphonie, op. 21. Journal of Music Theory, 1967, 11, 60 - 116

Hirsh, I.J., Auditory perception of temporal order. Journal of the Acoustical Society of America, 1959, 31, 759 - 767

Hirsh, I.J., Information processing in input channels for speech and language: The significance of serial order of stimuli. In: C.H. Millikan & F.L. Darley (Eds.), Brain mechanisms underlying speech and language. New York: Grune & Stratton, 1967

Hirsh, I.J., Temporal order and auditory perception. In: H.R. Moskowitz, B. Scharf & J.C. Stevens (Eds.), Sensation and measurement. Dordrecht: Reidel, 1974

Hirsh, I.J. & Fraisse, P., Simultanéité et succession de stimuli hétérogènes. Année de Psychologie, 1964, 64, 1 - 19

Hirsh, I.J. & Sherrick, C.E., Perceived order in different sense modalities. Journal of Experimental Psychology, 1961, 62, 423 - 432

Hitch, G.J., The role of attention in visual and auditory suffix effects. Memory & Cognition, 1975, 3, 501 - 505

Hochberg, J., In the mind's eye. In: R.N. Haber (Ed.), Contemporary theory and research in visual perception. New York: Holt, Rinehart & Winston, 1968

Hochberg, J., Organization and the Gestalt tradition. In: E.C. Carterette & M. Friedman (Eds.), Handbook of perception, vol. 1. New York: Academic Press, 1974

Hochberg, J. & Brooks, V., The psychophysics of form: Reversible perspective drawings of spatial objects. American Journal of Psychology, 1960, 73, 337 - 354

Hochberg, J. & McAlister, E., A quantitative approach to figural "goodness". Journal of Experimental Psychology, 1953, 46, 361 - 364

Hörmann, H., Meinen und Verstehen. Frankfurt: Surkamp, 1976

Hörmann, H., Psychologie der Sprache. Berlin - Heidelberg: Springer, 1977 (2. Aufl.

Holmes, V.M. & Foster, K.I., Detection of extraneous signals during sentence recognition. Perception & Psychophysics, 1970, 7, 297 - 301

Holmes, V.M. & Foster, K.I., Click localisation and syntactic structure. Perception & Psychophysics, 1972, 12, 9 - 15

ten Hoopen, G. & Vos, J., Attention-switching and grouping in counting interaurally presented clicks. Acta Psychologica, 1979, 43, 283 - 297

House, W.J., Octave generalisation and the identification of distorted melodies. Perception & Psychophysics, 1977, 21, 586 - 589

Hubel, D.H. & Wiesel, T.N., Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. Journal of Physiology, 1962, 160, 106 - 154

Hull, C.L., The concept of the habit-family hierarchy and maze learning. Psychological Review, 1934, 41, 33 - 54, 134 - 152

Huybrechts, R., Sequentielles Lernen und gedächtnismäßige Strukturbildungen. Zeitschrift für Psychologie, 1977, 185, 177 - 213, 375 - 396

Idson, W.L. & Massaro, D.W., Cross-octave masking of single tones and musical sequences: The effect of structure on auditory recognition. Perception & Psychophysics, 1976, 19, 155 - 175

Idson, W.L. & Massaro, D.W., A bidimensional model of pitch in the recognition of melodies. Perception & Psychophysics, 1978, 24, 551 - 565

James, W., Principles of psychology. New York: Dover, 1950 (Nachdruck der Ausgabe von 1890)

Johnson, N.F., Linguistic models and functional units of language behavior. In: S. Rosenberg (Ed.), Direction in psycholinguistics. New York: Macmillian, 1965

Johnson, N.F., Chunking: Associative chaining versus coding. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 725 - 731

Johnson, N.F., Chunking and organization in the process of recall. In: G.H. Bower (Ed.), The psychology of learning and motivation, vol. 4. New York: Academic Press, 1970

Johnson, N.F., Organization and the concept of a memory code. In: A.W. Melton & E. Martin (Eds.), Coding processes in human memory. Washington, D.C.: Winston, 1972

Johnson, N.F., Higher order encoding: Process or state. Memory & Cognition, 1973, 1, 491 - 494

Johnson, N.F., The memorial structure of organized sequences. Memory & Cognition, 1978, 6, 233 - 239 (a)

Johnson, N.F., Coding processes in memory. In: W.K. Estes (Ed.), Handbook of learning and cognitive processes, vol. 6. Hillsdale, N.J.: Erlbaum, 1978 (b)

Jones, M.R., Cognitive representations of serial patterns. In: B.H. Kantovitz (Ed.), Human information processing: Tutorials in performance and cognition. Hillsdale, N.J.: Erlbaum, 1974

Jones, M.R. Auditory patterns: Studies in the perception of structure. In: E.C. Carterette & M.P. Friedman (Eds.), Handbook of perception, vol. VIII. New York: Academic Press, 1978

Jones, M.R., Maser, D.J. & Kidd, G.R., Rate and structure in memory for auditory patterns. Memory & Cognition, 1978, 6, 246 - 258

Joos, M., Description of language design. Journal of the Acoustical Society of America, 1950, 22, 701 - 708

Kahneman, D., Attention and effort. Englewood Cliffs, N.J.: Prentice-Hall, 1973

Kallman, H.J. & Massaro, D.W., Tone chroma is functional in melody recognition. Perception & Psychophysics, 1979, 26, 32 - 36

Katz, D., Gestaltpsychologie. Basel: Schwabe, 1948

Keil, S., "Strukturelles Hören" - Versuch einer Begriffserklärung. Musik und Bildung, 1974, 6, 665 - 667 (a)

Keil, S., Zum Begriff 'Struktur' in der Musik. Musica, 1974, 28, 324 - 328 (b)

Kerivan, J.E. & Carey, B.J., Pattern identification of pure tones and frequency glides by untrained listeners. Perception & Psychophysics, 1976, 20, 489 - 492

Kinchla, R.A. & Wolfe, J.M., The order of visual processing: "Top-down", "bottom-up", or "middle-out". Perception & Psychophysics, 1979, 25, 225 - 231

Kintsch, W., The representation of meaning in memory. Hillsdale, N.J.: Erlbaum, 1974

Kleinen, G., Experimentelle Studien zum musikalischen Ausdruck. Dissertation, Hamburg, 1968

Kleinen, G., Entwicklungspsychologische Grundlagen musikalischen Verhaltens. In: H. Selger (Hg.), Musik und Musikunterricht in der Gesamtschule. Weinheim: Beltz, 1972

Klusen, E., Volkslieder aus 500 Jahren. Frankfurt: Fischer, 1978

Knaus, J. (Hg.), Sprache - Dichtung - Musik. Tübingen: Niemeyer, 1973

Kneif, T., Musik und Zeichen. Musica, 1973, 27, 9 - 12

Koch, W.A., Varia Semiotica. Hildesheim: Olms, 1971

Koestler, A., The act of creation. London: Hutchinson, 1964

Koffka, K., Experimental-Untersuchungen zur Lehre vom Rhythmus. Zeitschrift für Psychologie, 1909, 52, 1 - 109

Koffka, K., Principles of Gestalt Psychology. London: Routledge & Kegan Paul, 1935

Kratzer, A., Pause, E. & Stechow, A.v., Einführung in die Theorie und Anwendungen der generativen Syntax. Frankfurt: Athenäum, 1974

Kretzschar, H., Anregungen zur Förderung musikalischer Hermeneutik (1902). In: Gesammelte Aufsätze über Musik. Leipzig, 1910/11

Künding, A., Das Musikerlebnis in psychologischer und psychotherapeutischer Sicht mit besonderer Berücksichtigung seiner kompensatorischen Funktion. Winterthur, 1961

Kuhn, T.S., The structure of scientific revolutions. Chicago: University of Chicago Press, 1962

Kunst, J., Making sense in music. Ghent: University of Ghent, 1978

Kurth, E., Grundlagen des linearen Kontrapunkts. Bern, 1917

Kurth, E., Musikpsychologie. Berlin: Hesse, 1931

Ladefoged, P., Three areas of experimental phonetics. London: Oxford University Press, 1967

Ladefoged, P. & Broadbent, D.E., Perception of sequence in auditory events. Quarterly Journal of Experimental Psychology, 1960, 12, 162 - 170

Lashley, K.S., The problem of serial order in behavior. In: L.A. Jeffress (Ed), Cerebral mechanisms in behavior. New York: Wiley, 1951

Laske, O.E., The information processing approach to musical cognition. Interface, 1974, 3, 109 - 136

Laske, O.E., Toward a theory of musical cognition. Interface, 1975, 4, 147 - 208

Laske, O.E., Considering human memory in designing unser interfaces for computer music. Computer Music Journal, 1977, 2, 39 - 45 (a)

Laske, O.E., Music, memory, and thought. Ann Arbor: University Microfilms International, 1977 (b)

Laske, E., Toward an explicit cognitive theory of musical listening. Computer Music Journal, 1980, 4, 73 - 83

Leshowitz, B. & Hanzi, R., Serial position effects for tonal stimuli. Memory & Cognition, 1974, 2, 112 - 116

Lévi-Strauss, C., "Bolero" de Maurice Ravel. L'Homme, 1971, 11, 5 - 14

Lewin, D., Some applications of communication theory to the study of twelve-tone music. Journal of Music Theory, 1968, 12, 51 - 81

Lindblom, B. & Sundberg, J., Towards a generative theory of melody. Svensk Tidskrift för Musikforsking, 1970, 52, 71 - 88

Lipps, T., Zur Theorie der Melodie. Zeitschrift für Psychologie und Physiologie der Sinnesorgane, 1902, 27, 225 - 263

Lissa, Z., Ebenen des musikalischen Verstehens. In: Z.Lissa, Neue Aufsätze zur Musikästhetik. Wilhemshaven: Heinrichshofen, 1975

Lockhart, R.S., Craik, F.I.M. & Jacoby, L., Depth of processing, recognition, and recall. In: J. Brown (Ed.), Recall and recognition. New York: Wiley, 1976

Lundin, R.W., An objective psychology of music. New York: Ronald Press, 1967

Lupp, H., Hauske, G. & Wolf, W., Perceptual latencies to sinusoidal gratings. Vision Research, 1976, 16, 969 - 972

Lutz, W. (Hg.), Mein Kinderliederbuch. Mainz: Schott, 1950

Mace, W.M., Ecologically stimulating cognitive psychology: Gibsonian perspectives. In: W. Weimer & D. Palermo (Eds.), Cognition and the symbolic processes. Hillsdale, N.J.: Erlbaum, 1974

Mace, W.M., James J. Gibson's strategy for perceiving: Ask not what's inside your head, but what your head's inside of. In: R. Shaw & J.D. Bransford (Eds.), Perceiving, acting and knowing. Toward an ecological psychology. Hillsdale, N.J.: Erlbaum, 1977

Mach, E., Beiträge zur Analyse der Empfindungen. Jena: G. Fischer, 1886

Marks, L.E., The unity of the senses. Interrelations among the modalities. New York: Academic Press, 1978

Marmurek, H.H.C. & Johnson, N.F., Identification of words and letters within words. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, 1973 (zit. nach Johnson, 1973)

Marmurek, H.H.C. & Johnson, N.F., Hierarchical organization as a determinant of sequential learning. Memory & Cognition, 1978, 6, 240 - 245

Martin, J.G., Rhythmic (hierarchical) versus serial structure in speech and other behavior. Psychological Review, 1972, 79, 487 - 509

Maslow, A.H., Motivation and personality. New York: Harper, 1954

Massaro, D.W., Preperceptual auditory images. Journal of Experimental Psychology, 1970, 85, 411 - 417

Massaro, D.W., Preperceptual images, processing time, and perceptual units in auditory perception. Psychological Review, 1972, 79, 124 - 145

Massaro, D.W., Experimental psychology and information processing. Chicago: Rand McNally, 1975

Massaro, D.W., Auditory information processing. In: W.K. Estes (Ed.), Handbook of learning and cognitive processes, vol. 4. Hillsdale, N.J.: Erlbaum, 1976 (a)

Massaro, D.W., Perceiving and counting sounds. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 337 - 346 (b)

Massaro, D.W., Perceptual grouping in audition. Perception, 1977, 6, 541 - 553

Massaro, D.W. & Idson, W.L., Temporal course of perceived duration. Perception & Psychophysics, 1976, 20, 331 - 352

Massaro, D.W. & Idson, W.L., Backward recognition masking in relative pitch judgments. Perceptual and Motor Skills, 1977, 45, 87 - 97

Massaro, D.W. & Idson, W.L., Target-mask similarity in backward recognition masking of perceived tone duration. Perception & Psychophysics, 1978, 24, 225 - 236

McCawley, J.D., On what is deep about deep structures. In: W.B. Weimer & D.S. Palermo (Eds.), Cognition and the symbolic processes. Hillsdale, N.J.: Erlbaum, 1974

McClelland, J.L., Perception and masking of wholes and parts. Journal of Experimental Psychology: Human Perception and Performance, 1978, 4, 210 - 223

McDougall, R., The structure of simple rhythm forms. Psychological Review, Monograph Supplements, 1903, 4, 309 - 411

McNally, K.A. & Handel, S., Effect of element composition on streaming and the odering of repeating sequences. Journal of Experimental Psychology: Human Perception and Performance, 1977, 3, 451 - 460

McNeil, D., The acquisition of language: The study of developmental psycholinguistics. New York: Harper & Row, 1970

Meinong, A., Zur Psychologie der Komplexionen und Relationen. Zeitschrift für Psychologie, 1891, 2, 245 - 265

Meinong, A., Über Gegenstände höherer Ordnung und deren Verhältnis zur inneren Wahrnehmung. Zeitschrift für Psychologie, 1899, 21, 182 - 272

Metzger, W., Psychologie. Darmstadt: Steinkopff, 1963 (3. Aufl.)

Metzger, W., Figural-Wahrnehmung. In: W.Metzger & H.Erke (Hg.), Handbuch der Psychologie, Bd. 1, 1. Halbband. Göttingen: Hogrefe, 1966

Meumann, E., Untersuchungen zur Psychologie und Aesthetik des Rhythmus. Philosophische Studien, 1894, 10, 249 - 322, 393 - 430

Mewhort, D.J.K., Sequential redundancy and letter spacing as determinants of tachistoscopic recognition. Canadian Journal of Psychology, 1966, 20, 435 - 444

Mewhort, D.J.K., Scanning, chunking, and the familiarity effect in tachistoscopic recognition. Journal of Experimental Psychology, 1972, 93, 69 - 71

Mewhort, D.J.K., Accuracy and order of report in tachistoscopic identification. Canadian Journal of Psychology, 1974, 28, 383 - 398

Meyer, L.B., Emotion and meaning in music. Chicago: University of Chicago Press, 1956

Meyer, L.B., Meaning in music and information theory. Journal of Aesthetics and Art Criticism, 1957, 15, 412 - 424

Meyer, M., Elements of psychological theory of melody. Psychological Review, 1900, 7, 241 - 273

Michon, J.A., The making of the present: A tutorial review. In: J. Requin (Ed.) Attention and performance VII. Hillsdale, N.J.: Erlbaum, 1978

Miller, G.A., The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81 - 97 (a)

Miller, G.A., Human memory and the storage of information. IRE Transactions on Information Theory, 1956, IT-2, 129 - 137 (b)

Miller, G.A., Some psychological studies of grammar. American Psychologist, 1962, 17, 748 - 762

Miller, G.A., Bruner, J.S. & Postman, L., Familiarity of letter sequences and tachistoscopic identification. Journal of General Psychology, 1954, 50, 129 - 139

Miller, G.A., Galanter, E. & Pribram, K.H., Plans and the structure of behavior. New York: Holt, Rinehart & Winston, 1960

Miller, G.A. & Heise, G.A., The trill threshold. Journal of the Acoustical Society of America, 1950, 64, 637 - 638

Miller, G.A. & Isard, S., Some perceptual consequences of linguistic rules. In: L.A. Jakobovits & M.S. Miron (Eds.), Readings in the psychology of language. Englewood Cliffs, N.J.: Prentice-Hall, 1967

Minsky, M., A framework for representing knowledge. In: P.H. Winston (Ed.), The psychology of computer vision. New York: McGraw-Hill, 1975

Miyake, I., Researches on rhythmic action. Studies from the Yale Psychological Laboratory, 1902, 10, 1 - 48

Moles, A., Informationstheorie der Musik. Nachrichtentechnische Fachberichte, 1956, 3, 47 - 55

Moles, A., Informationstheorie und ästhetische Wahrnehmung. Köln: Du Mont, 1971

Moog, H., Zum Gegenstand der Musikpsychologie. Psychologische Rundschau, 1977, 28, 110 - 125

Moray, N., Attention: Selective processes in vision and hearing. London: Hutchinson, 1969

Moray, N., Bates, A. & Barnett, I., Experiments on the four-eared man. Journal of the Acoustical Society of America, 1965, 38, 196 - 201

Morrison, D.F., Multivariate statistical methods. New York: McGraw-Hill, 1967

Morton, J., Crowder, R.G. & Prussin, H.A., Experiments with the stimulus suffix effect. Journal of Experimental Psychology Monograph, 1971, 91, 169 - 190

de la Motte-Haber, H., Musikpsychologie. Köln: Gerig, 1972

de la Motte-Haber, H., Psychologie und Musiktheorie. Frankfurt: Diesterweg, 1976

Mull, H.K., Prefered regions in music compositions and the effect of repetition upon them. American Journal of Psychology, 1940, 53, 583 - 586

Mull, H.K., The effect of repetition upon the enjoyment of modern music. Journal of Psychology, 1957, 43, 155 - 162

Murdock, B.B., The retention of individual items. Journal of Experimental Psychology, 1961, 62, 618 - 625

Mursell, J.L.,The psychology of music. New York: Norton, 1937

Myers, C.S., Individual differences in the listening to music. British Journal of Psychology, 1922, 13, 52 - 71

Myers, C.S. & Valentine, C.W., A study of individual differences in attitude toward tones. British Journal of Psychology, 1914, 7, 68 - 111

Nadel, S., Der duale Sinn der Musik. Regensburg, 1931

Nattiez, J.J., Is a descriptive semiotics of music possible? Language Sciences, 1972, 23, 1 - 7

Nattiez, J.J., Linguistics: A new approach for musical analysis? International Review of the Aesthetics and Sociology of Music, 1973, 4, 51 - 69

Navon, D., Forest before trees: The precedence of global features in visual perception. Cognitive Psychology, 1977, 9, 353 - 383

Neisser, U., Cognitive psychology. New York: Appleton-Century-Crofts, 1967

Neisser, U., Cognition and reality. San Francisco: Freeman, 1976

Neisser, U. & Hirst, W., Effect of practice on the identification of auditory sequences. Perception & Psychophysics, 1974, 15, 391 - 398

Nettl, B., Some linguistic approaches to musical analysis. Journal of the International Folk Music Council, 1958, 10, 37 - 41

Neumann, O., Aufmerksamkeit als 'zentrale Verarbeitungskapazität'. Anmerkungen zu einer Metapher. In: M. Tücke & W. Deffner (Hg.), Bericht über die 2. Osnabrücker Arbeitstagung Psychologie, Borkum, 1978. Psychologische Forschungsberichte aus dem Fachbereich 3 der Universität Osnabrück. Nr. 10, 1978

Neumann, O., Informatiosselektion und Handlungssteuerung. Untersuchungen zur Funktionsgrundlage des Stroop-Interferenzphänomens. Dissertation Bochum, 1980

Nickerson, R.S. & Freeman, B., Discrimination of order of the components of repeating tone sequences: Effects of frequency separation and extensive practice. Perception & Psychophysics, 1974, 16, 471 - 477

van Noorden, L.P.A.S., Temporal coherence in the perception of tone sequences. Dissertation Technische Universität Eindhoven, 1975

van Noorden, L.P.A.S., Minimum differences of level and frequency for perceptual fission of tone sequences ABAB. Journal of the Acoustical Society of America, 1977, 61, 1041 - 1045

Nordmark, J.O., Frequency and periodicity analysis. In: E.C. Carterette & M.P. Friedman (Eds.), Handbook of perception, vol. IV. New York: Academic Press, 1978

Norman, D.A. & Rumelhart, D.E. (Eds.), Explorations in cognition. San Francisco: Freeman, 1975

Olshowky, T., On the notion of a rule. In: T. Storer & D. Winter (Eds.), Formal aspects of cognitive processes. Berlin: Springer, 1975

Ornstein, R.E., On the experience of time. Harmondsworth: Penguin Books, 1969

Ortmann, O., On the melodic relativity of tones. Psychological Monographs, 1926, 35, No. 162

Ortmann, O., Some tonal determinants of melodic memory. Journal of Educational Psychology, 1933, 24, 454 - 467

Osmond-Smith, D., Problems of terminology and method in the semiotics of music. Semiotica, 1974, 11, 269 - 294

Palmer, S.E., Visual perception and world knowledge. In: D.A. Norman & D.E. Rumelhart (Eds.), Explorations in cognition. San Francisco: Freeman, 1975

Palmer, S.E., Hierarchical structure in perceptual representation. Cognitive Psychology, 1977, 9, 441 - 474

Palmer, S.E., Fundamental aspects of cognitive representation. In: E.Rosch & B.B. Lloyd (Eds.), Cognition and categorization. Hillsdale, N.J.: Erlbaum, 1978 (a)

Palmer, S.E., Structural aspects of visual similarity. Memory & Cognition, 1978, 6, 91 - 97 (b)

Patterson, J.H. & Green, D.M., Discrimination of signals having identical energy spectra. Journal of the Acoustical Society of America, 1970, 48, 894 - 905

Patterson, R.D. & Green, D.M., Auditory masking. In: E.C. Carterette & M.P. Friedman (Eds.), Handbook of perception, vol. IV. New York: Academic Press, 1978

Perkins, D.N., Coding position in a sequence by rhythmic groupings. Memory & Cognition, 1974, 2, 219 - 223

Peters, R., Perceived order of tone pulses. Journal of the Acoustical Society of America, 1964, 36, 1042

Peters, R., Perceived order of tone pulses. Journal of the Acoustical Society of America, 1967, 42, 1216

Peters, R., The measurement of temporal factors in auditory perception. In: P. Singh (Ed.), Measurement procedures in speech, hearing, and language. Baltimore: University Park Press, 1975

Peters, R. & Wood, T.J., Perception of temporal order for tones. Journal of the Acoustical Society of America, 1973, 53, 311 - 312 (a)

Peters, R. & Wood, T.J., Perceived order of tone pulses. Journal of the Acoustical Society of America, 1973, 54, 315 (b)

Petri, H., Form- und Strukturparallelen in Literatur und Musik. Studium Generale, 1966, 19, 72 - 84

Piaget, J., The origins of intelligence in children. New York: International Universities Press, 1952

Pinheiro, M.L. & Ptacek, P.H., Reversals in the perception of noise and tone patterns. Journal of the Acoustical Society of America, 1971, 49, 1778 - 1782

Pinkerton, R.C., Information theory and melody. Scientific American, 1956, 194, 77 - 86

Polanyi, M., Tacit knowing: Its bearing on some problems of philosophy. Review of Modern Physics, 1962, 34, 601 - 616

Polanyi, M., The tacit dimension. Garden City: Doubleday, 1966

Pollack, I. & Johnson, L.B., Memoryspan with efficient coding procedures. American Journal of Psychology, 1965, 78, 609 - 614

Pomerantz, J.R., Are complex visual features derived from simple ones? In: E. Leeuwenberg & H. Buffart (Eds.), Formal theories of visual perception. Sussex: Wiley, 1978

Pomerantz, J.R. & Garner, W.R., Stimulus configuration in selective attention tasks. Perception & Psychophysics, 1973, 14, 565 - 569

Pomerantz, J.R. & Sager, L.C., Line-slope vs. line-arrangement discrimination: A comment on Ambler & Finklea's paper. Perception & Psychophysics, 1976, 20, 220

Pomerantz, J.R. & Schaïtzberg, S.D., Grouping by proximity: Selective attention measures. Perception & Psychophysics, 1975, 18, 355 - 361

Posner, M.I., Immediate memory in sequential tasks. Psychological Bulletin, 1963, 60, 333 - 349

Pratt, C.C., Music as the language of emotion. Washington: The Library of Congress, 1952

Preusser, D., The effect of structure and rate on the recognition and description of auditory temporal patterns. Perception & Psychophysics, 1972, 11, 233 - 240

Preusser, D., Garner, W.R. & Gottwald, R.L., Perceptual organization of two-element temporal patterns as a function of their component one-element patterns. American Journal of Psychology, 1970, 83, 151 - 170 (a)

Preusser,D., Garner, W.R. & Gottwald, R.L., The effect of starting pattern on descriptions of perceived temporal patterns. Psychonomic Science, 1970, 21, 219 - 220 (b)

Pribram, K.H., The nature of the perceived universe. In: R. Shaw & J.D. Bransford (Eds.), Perceiving, acting, and knowing. Toward an ecological psychology. Hillsdale, N.J.: Erlbaum, 1977 (a)

Pribram, K.H., Holonomy and the structurein the organization of perception. In: J.M. Nicholas (Ed.), Images, perception, and knowledge. Dordrecht: Reidel, 1977 (b)

Pribram, K.H., Nuwer, M. & Baron, R., The holographic hypothesis of memory structure in brain function and perception. In: R.C. Atkinson, D.H. Krantz, R.C. Luce & P. Suppes (Eds.), Contemporary developments in mathematical psychology. San Francisco: Freeman, 1974

Prinzmetal, W. & Banks, W.P., Good continuation affects visual detection. Perception & Psychophysics, 1977, 21, 389 - 395

Ptacek, P.H. & Pinheiro, M.L., Pattern reversal in auditory perception. Journal of the Acoustical Society of America, 1971, 49, 493 - 498

Quillian, M.R., Semantic memory. In: M. Minsky (Ed.), Semantic information processing. Cambridge, Mass.: MIT-Press, 1968

Raab, D.H., Backward masking. Psychological Bulletin, 1963, 60, 118 - 129

Rauhe, H., Kritischer Schallplattenvergleich aus den Bereichen Folklore und Beat. In: E. Kraus (Hg.), Der Einfluß der technischen Mittler auf die Musikerziehung unserer Zeit. Mainz: Schott, 1968

Rauhe, H., Reinecke, H.-P. & Ribke, W., Hören und Verstehen. München: Kösel, 1975

Reber, A.S., Implicit learning of artificial grammars. Journal of Verbal Learning and Verbal Behavior, 1967, 6, 855 - 863

Reber, A.S., Transfer of syntactic structure in synthetic languages. Journal of Experimental Psychology, 1969, 81, 115 - 119

Reber, A.S., Locating clicks in sentences: Left, center, and right. Perception & Psychophysics, 1973, 13, 133 - 138 (a)

Reber, A.S., What clicks may tell us about speech perception. Journal of Psycholinguistic Research, 1973, 2, 286 (b)

Reber, A.S., Implicit learning of synthetic languages: The role of instructional set. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 88 - 94

Reber, A.S. & Anderson, J.R., The perception of clicks in linguistic and nonlinguistic messages. Perception & Psychophysics, 1970, 8, 81 - 89

Reber, A.S. & Lewis, S., Implicit learning: An analysis of the form and structure of a body of tacit knowledge. Cognition, 1977, 5, 333 - 361

Reed, S.K., Schemes and theories of pattern recognition. In: E.C. Carterette & M.P. Friedman (Eds.), Handbook of perception, vol. IX. New York: Academic Press, 1978

Reinert, G., Visuelles Suchen in Strichzeichnungen unterschiedlicher Organisation bei Variation der situativen Nähe der Suchobjekte. Bericht Nr. 13, Psychologisches Institut der Ruhr-Universität Bochum, Arbeitseinheit Kognitionspsychologie, 1980

Restle, F., Theory of serial pattern learning: Structural trees. Psychological Review, 1970, 77, 481 - 495

Restle, F., Critique of pure memory. In: R.L. Solso (Ed.), Theories in cognitive psychology: The Loyola symposium. Potomac, Md.: Erlbaum, 1974

Restle, F. & Brown, E., Organization of serial pattern learning. In: G.H. Bower (Ed.), The psychology of learning and motivation, vol. 4. New York: Academic Press, 1970

Révész, G., Einführung in die Musikpsychologie. Bern, 1946

Riemann, H., Musikalische Formenlehre, 1889 (Grundriß der Kompositionslehre. Berlin, 1922, 9. Aufl.)

Riemann, H., Ideen zu einer Lehre von den Tonvorstellungen. Jahrbuch der Musikbibliothek Peters 21/22. Leipzig: Peters, 1914/15

Rigg, M., The mood effects of music. A comparison of data from four different investigations. Journal of Psychology, 1964, 58, 427 - 438

Risset, J.C., Musical acoustics. In: E.C. Carterette & M.P. Friedman (Eds.), Handbook of perception, vol. IV. New York: Academic Press, 1978

Risset, J.C. & Matthews, M.V., Analysis of musical instrument tones. Physics Today, 1969, 22, 23 - 30

Rosch, E., Principles of categorization. In: E. Rosch & B.B. Lloyd (Eds.), Cognition and categorization. Hillsdale, N.J.: Erlbaum, 1978

Royer, F.L. & Garner, W.R., Response uncertainty and perceptual difficulty of auditory temporal patterns. Perception & Psychophysics, 1966, 1, 41 - 47

Royer, F.L. & Garner, W.R., Perceptual organization of nine-element auditory temporal patterns. Perception & Psychophysics, 1970, 7, 115 - 120

Rumelhart, D.E., A multicomponent theory of perception of briefly exposed visual displays. Journal of Mathematical Psychology, 1970, 7, 191 - 218

Rumelhart, D.E., Notes on a schema for stories. In D.G. Bobrow & A. Collins (Eds.), Representation and understanding. New York: Academic Press, 1975

Ruwet, N., Language, Musique, Poésie. Paris: Le Seuil, 1972

Sanders, A.F., Psychologie der Informationsverarbeitung. Stuttgart: Huber, 1971

Sanders, D.A., Auditory perception of speech. An introduction to principles and problems. Englewood Cliffs, N.J.: Prentice-Hall, 1977

Sapir, J.D., Diola-Togny funeral songs and the naive critic. African Language Review, 1969, 8, 176 - 191

Saporta, S., Psycholinguistic theory of generative grammar. Eugene, Origon: State System of Higher Education, 1967

Schaff, A., Das Verstehen der verbalen Sprache und das "Verstehen" von Musik. In: P. Faltin & H.-P. Reinecke (Hg.), Musik und Verstehen. Köln: Volk, 1973

Scharf, B., Loudness. In: E.C. Carterette & M.P. Friedman (Eds.), Handbook of perception, vol. IV. New York: Academic Press, 1978

Schenker, H., Harmony. Chicago: University of Chicago Press, 1954

Schering, A., Zur Sinndeutung der 4. und 5. Symphonie von Beethoven. Zeitschrift für Musikwissenschaft, 1934, 16, 65 - 83

Schoen, M. (Ed.), The effects of music. New York: Hartcourt, Brace & World, 1927

Schole, H., Tonpsychologie und Musikästhetik. Göttingen, 1930

Schüngeler, H. (Hg.), Das neue Volksliederbuch. Mainz: Schott, o.J.

Seashore, C.E., Psychology of music. New York: McGraw, 1938

Seitz, M.R. & Weber, B.A., Effects of response requirements on the location of clicks superimposed on sentences. Memory & Cognition, 1974, 2, 43 - 46

Selfridge, O.G., Pandemonium: A paradigm for learning. In: The mechanization of thought processes. London: H.M. Stationery Office, 1959

Shannon, C.E. Weaver, W., The mathematical theory of communication. Urbana: University of Illinois Press, 1949

Shapere, D., Scientific theories and their domains. In: F. Suppe (Ed.), The structure of scientific theories. Urbana, Ill.: University of Illinois Press, 1974

Shatin, L., Alternation of mood via music: A study of the vectoring effect. Journal of Psychology, 1970, 75, 81 - 86

Shaw, A.C., Picture grammars and parsing. In: G.J. Dalenoort (Ed.), Process models for psychology. Rotterdam: Rotterdam University Press, 1973

Shaw, R. & Bransford, J., Introduction: Psychological approaches to the problem of knowledge. In: R. Shaw & J.D. Bransford (Eds.), Perceiving, acting, and knowing. Toward an ecological psychology. Hillsdale, N.J.: Erlbaum, 1977

Shaw, R. & McIntyre, M., Algoristic foundations to cognitive psychology. In: W.B. Weimer & D.S. Palermo (Eds.), Cognition and the symbolic processes. Hillsdale, N.J.: Erlbaum, 1974

Shaw, R., McIntyre, M. & Mace, W., The role of symmetry in event perception. In: R.B. MacLeod & H.L. Pick (Eds.), Perception. Essays in honor of James J. Gibson. Ithaca and London: Cornell University Press, 1974

Shaw, R. & Pittenger, J., Perceiving change. In: H.L. Pick & E. Saltzman (Eds.), Models of perceiving and processing information. Hillsdale, N.J.: Erlbaum, 1978

Shuter, R., The psychology of musical ability. London: Methuen, 1968

Siegel, W., Memory effects in the method of absolute judgment. Journal of Experimental Psychology, 1972, 94, 121 - 131

Siegel, J.A. & Siegel, W., Absolute identification of notes and intervals by musicans. Perception & Psychophysics, 1977, 21, 143 - 152 (a)

Siegel, J.A. & Siegel, W., Categorial perception of tonal intervals: Musicans can't tell s h a r p from f l a t . Perception & Psychophysics, 1977, 21, 399 - 407 (b)

Simon, H.A., Induction and representation of sequential patterns. In: E.L.J. Leeuwenberg & H.F.J.M. Buffart (Eds.), Formal theories of visual perception. New York: Wiley, 1978

Simon, H.A. & Sumner, R.K., Pattern in music. In: B. Kleinmutz (Ed.), Formal representations of human judgment. New York: Wiley, 1968

Sloboda, J.A., The eye-hand span: An approach to the study of sight-reading. Psychology of Music, 1974, 2, 4 - 10

Sloboda, J.A., Visual perception of musical notation: Registering pitch symbols in memory. Quarterly Journal of Experimental Psychology, 1976, 28, 1 - 16

Sloboda, J.A., Phrase units as determinants of visual processing in music reading. British Journal of Psychology, 1977, 68, 117 - 124

Smoliar, S.W., A parallel processing model of musical structures. Cambridge, Mass.: MIT, 1971

Sperling, G., The information available in brief visual presentations. Psychological Monographs: General and Applied, 1960, 74, No. 498

Sperling, G.A., A model for visual memory tasks. Human Factors, 5, 19 - 31

Spieth, W., Curtis, J.F. & Webster, J.C., Responding to one of two simultaneous messages. Journal of the Acoustical Society of America, 1954, 26, 391 - 396

Spoehr, K.T. & Smith, E.E., The role of orthographic and phonotactic rules in perceiving letter patterns. Journal of Experimental Psychology: Human Perception and Performance, 1975, 1, 21 - 34

Springer, G.P., Language and music: Parallels and divergencies. In: M. Halle, H.G. Lunt, H. McLean & C.H. Van Schooneveld (Eds.), For Roman Jakobson. Den Haag: Mouton, 1956

Stahl, V., Informationswissenschaft und Musikanalyse. Grundlagenstudien aus Kybernetik und Geisteswissenschaft, 1964, 5, 51 - 58

Sternberg, S., High-speed scanning in human memory. Science, 1966, 153, 652 - 654

Sternberg, S., The discovering of processing stages: Extension of Donders' method. Acta Psychologica, 1969, 30, 276 - 315

Stetson, R.H., A motor theory of rhythm and discrete succession. Psychological Review, 1905, 12, 250 - 270, 292 - 350

Stevens, S.S. & Davis, H., Hearing. New York: Wiley, 1938

Stoffer, T.H., Aspekte einer generativen Syntax zur Beschreibung musikalischer Strukturen für eine kognitive Musikpsychologie. Bericht Nr. 11, Psychologisches Institut der Ruhr-Universität Bochum, Arbeitseinheit Kognitionspsychologie, 1979

Stoffer, T.H. & Boroda, M.G., Mathemathisierungstendenzen in der Musikwissenschaft: Eine Bibliographie zur quantitativ-statistischen und algebraisch-formalen Analyse musikalischer Strukturen. In: W. Matthäus (Hg.), Glottometrika 3. Quantitative Linguistics, vol. 5. Bochum: Brockmeyer, 1980

Stroop, J.R., Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 1935, 18, 643 - 662

Stumpf, C., Tonpsychologie. Leipzig: Hirzel, 1883/90 (2 Bände)

Sturges, P.T. & Martin, J.G., Rhythmic structure in auditory temporal pattern perception and immediate memory. Journal of Experimental Psychology, 1974, 102, 377 - 383

Suga, N., Functional properties of auditory neurons in the cortex of echo-locating bats. Journal of Physiology, 1965, 181, 671 - 700

Sundberg, J. & Lindblom, B., Generative theories in language and music descriptions. Cognition, 1976, 9, 99 - 122

Sutherland, N.S., Stimulus analyzing mechanisms. In: The mechanization of thought processes. London: H.M. Stationary Office, 1959

Sutherland, N.S., Intelligent picture processing. Paper presented at the Conference on the Evolution of the Nervous System and Behavior, Florida State University, Tellahassee, 1973 (zit. nach Turvey, 1975)

Terhardt, E., Psychoacoustic evaluation of musical sounds. Perception & Psychophysics, 1978, 23, 483 - 492

Thomas, I.B. & Fitzgibbons, P.J., Temporal order and perceptual class. Journal of the Acoustical Society of America, 1971, 50, 86

Thompson, A.L. & Klatzky, R.L., Studies of visual synthesis: Integration of fragments into forms. Journal of Experimental Psychology: Human Perception and Performance, 1978, 4, 244 - 263

Thorndyke, P.W., Cognitive structures in comprehension and memory of narrative discourse. Cognitive Psychology, 1977, 9, 77 - 110

Titchener, E.B., Lectures on the experimental psychology of the thought processes. New York: Macmillan, 1909

Tobias, J.V. (Ed.), Foundations of modern auditory theory, vol. 1. New York: Academic Press, 1970

Toda, M., Time and the structure of human cognition. In: J.T. Fraser & N. Lawrence (Eds.), The study of time. New York: Springer, 1975

Toppino, T.C., The underlying structures of sentences are not the primary units of speech processing: A reinterpretation of Bever, Lackner and Kirk's findings. Perception & Psychophysics, 1974, 15, 517 - 518

Treisman, A., Selective attention in man. British Medical Bulletin, 1964, 20, 12 - 16

Treisman, M. & Rostron, A.B., Brief auditory storage: a modification of Sperling's paradigm applied to audition. Acta Psycholgica, 1972, 36, 161 - 170

Turvey, M.T., On peripheral and central processes in vision: inferences from an information-processing analysis of masking with patterned stimuli. Psychological Review, 1973, 80, 1 - 52

Turvey, M.T., Perspectives in vision: Conception or perception? In: D.D. Duane & M.B. Rawson (Eds.), Reading, perception, and language. Baltimore: York Press, 1975

Turvey, M.T., Preliminaries to a theory of action with reference to vision. In: R.Shaw & J. Bransford (Eds.), Perceiving, acting, and knowing. Toward an ecological psychology. Hillsdale, N.J.: Erlbaum, 1977

Tversky, A., Features of similarity. Psychological Review, 1977, 84, 327 - 352

Uhr, L., "Pattern recognition" computers as models for form perception. Psychological Bulletin, 1963, 60, 40 - 73

Uttal, W.R., An autocorrelation theory of form detection. Hillsdale, N.J.: Erlbaum, 1975

Vernon, P.E., Auditory perception I: The Gestalt approach. British Journal of Psychology, 1934, 25, 123 - 139

Verveer, E.M., Barry, H. & Bousfield, W.A., Change in affectivity with repetition. American Journal of Psychology, 1933, 45, 130 - 134

Vitz, P.C., Affect as a function of stimulus variation. Journal of Experimental Psychology, 1966, 71, 74 - 79

Vitz, P.C., Experiments of affective reactions to 'music' stimuli. Sciences de l'Art - Scientific Aesthetics, 1974, 9, 3 - 14

Ward, W.D., Musical perception. In: J.V. Tobias (Ed.), Foundations of modern auditory theory, vol. 1. New York: Academic Press, 1970

Warren, R.M., Perception of temporal order: Special rules for the initial and terminal sounds. Journal of the Acoustical Society of America, 1972, 52, 167

Warren, R.M., Auditory temporal discrimination by trained listeners. Cognitive Psychology, 1974, 6, 237 - 256 (a)

Warren, R.M., Auditory pattern recognition by untrained listeners. Perception & Psychophysics, 1974, 15, 495 - 500 (b)

Warren, R.M. & Ackroff, J.M., Two types of auditory sequence perception. Perception & Psychophysics, 1976, 20, 387 - 394

Warren, R.M. & Byrnes, D.L., Temporal discrimination of recycled tonal sequences: Pattern matching and naming of order by untrained listeners. Perception & Psychophysics, 1975, 18, 273 - 280

Warren, R.M., Obusek, C.J., Farmer, R.M. & Warren, R.P., Auditory sequence: Confusion of patterns other than speech or music. Science, 1969, 164, 586 - 587

Watson, C.S., Wroton, H.W., Kelly, W.J. & Benbassat, C.A., Factors in the discrimination of tone patterns: I. Component frequency, temporal position, and silent intervals. Journal of the Acoustical Society of America, 1975, 57, 1175 - 1185

Watt, W.C., Autonome Linguistik, Psycholinguistik, kognitive Linguistik. In: H. Leuninger, M.H. Miller & F. Müller (Hg.), Linguistik und Psychologie, Bd. 1. Frankfurt: Athenäum, 1973

Weisstein, N.& Harris, C.S., Visual detection of line segments: An object superiority effect. Science, 1974, 186, 752 - 755

Wellek, A., Musikpsychologie und Musikästhetik. Frankfurt: Akademische Verlagsgesellschaft, 1963

Werbig, H., Informationsgehalt und emotionale Wirkungen von Musik. Mainz: Schott, 1971

Werner, H., Rhythmik, eine mehrwertige Gestaltverkettung. Zeitschrift für Psychologie, 1919, 82, 198 - 218

Werner, H., Über Mikromelodik und Mikroharmonik. Zeitschrift für Psychologie, 1926, 98, 74 - 89

Werner, H., Musical "micro-scales" and "micro-melodies". Journal of Psychology, 1940, 10, 149 - 156

White, B.W., Recognition of distorted melodies. American Journal of Psychology, 1960, 73, 100 - 107

Whitfield, I.C. & Evans, E.F., Responses of auditory cortical neurons to stimuli of changing frequency. Journal of Neurophysiology, 1965, 28, 655 - 672

Wickelgren, W.A., Consolidation and retroactive interference in short-term recognition memory for pitch. Journal of Experimental Psychology, 1966, 72, 250 - 259

Wickelgren, W.A., Auditory or articulatory coding in verbal short-term memory. Psychological Review, 1969, 76, 232 - 235

Williams, A. & Weisstein, N., Line segments are perceived better in a coherent context then alone: An object-line effect in visual perception. Memory & Cognition, 1978, 6, 85 - 90

Winckel, F., Die informationstheoretische Analyse musikalischer Strukturen. Musikforschung, 1964, 17, 1 - 14

Winer, B.J., Statistical principles in experimental design. New York: McGraw-Hill, 1970

Winograd, T., Linguistics and computer analysis of tonal harmony. Journal of Music Theory, 1968, 12, 2 - 49

Winograd, T., Understanding natural language. Edinburgh: Edinburgh University Press, 1972

Winston, P.H. (Ed.), The psychology of computer vision. New York: McGraw-Hill, 1975

Winzenz, D., Group structure and coding in serial learning. Journal of Experimental Psychology, 1972, 92, 8 - 19

Wirth, W., Die unmittelbare Teilung einer gegebenen Zeitstrecke. American Journal of Psychology, 1937, 50, 79 - 96

Witte, W., Zur Geschichte des psychologischen Ganzheits- und Gestaltbegriffs. Studium Generale, 1952, 5, 455 - 464

Wolff, P., Mechanismus der Redundanzausnutzung bei der Identifikation tachistoskopischer Reize. Diplomarbeit, Psychologisches Institut der Ruhr-Universität Bochum, 1971

Wolff, P., Entnahme der Identitäts- und Positionsinformation bei der Identifikation tachistoskopischer Buchstabenzeilen. Ein theoretischer und experimenteller Beitrag zur Grundlagenforschung des Lesens. Dissertation, Bochum, 1977

Woodrow, H.S., Time perception. In: S.S. Stevens (Ed.), Handbook of experimental psychology. New York: Wiley, 1951

Wundt, W., Grundzüge der physiologischen Psychologie. Leipzig: Engelmann, 1874

Youngblood, J.E., Style as information. Journal of Music Theory, 1958, 2, 24 - 35

Ziegenrücker, W., Allgemeine Musiklehre. München: Goldmann/Schott, 1979

Zwislocki, J., Theory of temporal auditory summation. Journal of the Acoustical Society of America, 1960, 32, 1046 - 1060

Zwislocki, J.J., Masking: Experimental and theoretical aspects of simultaneous forward, backward, and central masking. In: E.C. Carterette & M.P. Friedman (Eds.), Handbook of perception, vol. IV. New York: Academic Press, 1978

ANHANG 1

Syntax der in den Experimenten verwandten Melodien

1. Eine generative kontextabhängige Transformationssyntax einsätziger deutscher Volks- und Kinderlieder

Wir wollen an einem bestimmten musikalischen Formtypus zeigen, wie man zu einer streng formalen Strukturbeschreibung unter Verwendung linguistischer Methoden kommen kann.

1.1 Methode der syntaktischen Musikanalyse

1.1.1 Der Corpus

Will man künstlich generierte Musik in musikpsychologischen Untersuchungen verwenden, dann sollte man einen Formtypus wählen, der 1. den potentiellen Versuchspersonen möglichst gut bekannt ist, 2. von dem es sehr viele Exemplare gibt und der 3. eine möglichst regelmäßige Struktur aufweist. Bekannt sollte dieser Formtypus sein, weil es wichtige musikpsychologische Fragestellungen gibt, die eine Einbeziehung des Wissens des Hörers über die musikalische Struktur in ein Modell des Musikhörens erfordern, viele Exemplare dieses Formtypus' sollte es geben, weil eine empirische Analyse des zu diesem Formtypus gehörenden musikalischen Materials um so besser verallgemeinerungsfähig ist, je größer die zugrundeliegende Datenbasis ist, und eine möglichst große Regelmäßigkeit der Strukturen ist deshalb wünschenswert, weil diese 1. die syntaktische Beschreibung vereinfacht, aber 2. auch die Chance vergrößert, syntaxabhängige Effekte bei kognitiven Prozessen zu entdecken.

Aus diesen Gründen haben wir den Formtypus des Liedes gewählt. Innerhalb der Klasse der deutschen Volks- und Kinderlieder haben wir uns auf den einsätzigen Liedtypus beschränkt, der nur aus einer Periode besteht. Die für die Liedperiode gefundenen Regelmäßigkeiten lassen sich dann ohne größere Modifikationen an den Regeln der Syntax durch Einführung einer Regel verallgemeinern, die Verknüpfungen von Perioden erlaubt.

Der von uns analysierte Corpus enthält 71 einsätzige Volks- und Kinderlieder des 17. bis 20. Jahrhunderts. Als Quelle dienten uns LUTZ (1950), SCHÜNGELER (o.J.) und KLUSEN (1978). Titel und Quellen der Lieder sind im Anhang zusammengestellt.

1.1.2 Distributionsanalysen

Die Distributionsanalyse ist ein empirisches Verfahren zur Bestimmung der syntagmatischen und paradigmatischen Beziehungen zwischen den syntaktisch zu beschreibenden Elementen einer Struktur. Syntagmatische Beziehungen sind Beziehungen des Miteinandervorkommens von Einheiten, paradigmatische Beziehungen bestehen zwischen den Elementen, die an gleicher Position erscheinen können. Die syntagmatischen Beziehungen machen daher die Phrasenstruktur innerhalb einer musikalischen Periode aus, während die paradigmatischen Beziehungen nur festlegen, wie die abstrakte Periodenstruktur, etwa als leeres Muster von Positionen betrachtet, in der konkreten Melodie ausgefüllt werden kann.

Die stufenweise Teilung einer Periode heißt analog zur linguistischen Satzanalyse IC-Analyse, d.h. Analyse in "immediate constituents". Die Elementfolgen, die der Teilung unterworfen werden, heißen Konstitute, die Teile eines Konstituts sind Konstituenten. Die Teile, die durch eine einmalige Teilung eines Konstituts entstehen, sind die unmittelbaren Konstituenten dieses Konstituts.

Konstituenten können weiter geteilt werden, wodurch sie zu Konstituten werden, die ihrerseits in die unmittelbaren Konstituenten geteilt werden können.

Eine Periode kann nur geteilt werden, indem sie mit anderen Perioden verglichen wird. Ein Teilungsschritt ist nur dann gerechtfertigt, wenn es andere Konstitute gibt, die die gleichen Konstituenten in anderer Umgebung enthalten. Die IC-Analyse verweist auf die strukturellen Muster, nach denen Perioden gebildet werden. Um das System zu charakterisieren, reicht der konkrete, durch Teilung gewonnene Stammbaum jedoch nicht aus, weshalb man an den Knoten des Stammbaumes anstelle der Konstituenten die Bezeichnungen der Konstituentenklassen (Paradigmen) angibt, von denen sie Elemente sind.

Um die Zugehörigkeit einer Elementenfolge zu einer Konstituentenklasse festzustellen und damit eine Teilung zu rechtfertigen, werden Distributionsanalysen vorgenommen. Da eine Teilung nur dann sinnvoll ist, wenn sie möglichst viele Elemente umfaßt, die im gleichen Kontext vorkommen, gilt es den Grenzen, die mögliche Konstituentenklassen auszeichnen, besondere Aufmerksamkeit zu schenken. Eine Teilung der Konstituentenklasse "Periode" (Symbol P) in ihre unmittelbaren Konstituenten ist an den ausgezeichneten Stellen möglich, die den stärksten Restriktionen unterliegen. Die musikalische Distributionsanalyse besteht im wesentlichen im Auffinden derartiger Stellen sowie in der Analyse und Aufzählung der Kombinationsmöglichkeiten von Konstituentenklassen innerhalb einer Periode.

Wir haben das Auftreten der einzelnen Tonstufen der diatonischen Tonleiter an den Positionen innerhalb einer Periode sowie deren harmonische Funktion im Kontext der Nachbarpositionen untersucht. Außerdem wurden Tondaueranalysen, Pausenanalysen sowie Analysen der ordinalen Transformationsbeziehungen auf allen Teilungsebenen der Konstituentenhierarchie vorgenommen. Eine ausführliche tabellarische Darstellung der Ergebnisse der Distributionsanalysen

ist hier aus Platzgründen nicht möglich, jedoch geht aus den Regeln der Syntax implizit hervor, welche Analysen mit welchem Ergebnis vorgenommen wurden, damit die jeweilige Regel in der dargestellten Form formalisiert werden konnte, so daß der Leser im Prinzip durch eine eigene Distributionsanalyse des im Anhang angegebenen musikalischen Materials das Zutreffen der Regeln objektiv überprüfen kann.

1.1.3 Konstituentenstrukturen der untersuchten Lieder

Im folgenden geben wir eine Darstellung der im untersuchten Material realisierten Konstituentenstrukturen mit Angabe der distributiven Merkmale, die die Grenzen der Konstituentenklassen auszeichnen. Abb. 1 stellt einen Teil des sich aus den Distributionsanalysen ergebenden Strukturbaums dar, in dem optionale Konstituentenklassen eingeklammert sind. Die durch ordinale Transformationen ermöglichten Erweiterungen sind nicht enthalten; diese können aus den Regeln selbst entnommen werden. Die fehlenden Teile sind analog den dargestellten aufgebaut (s.u.).

Die Periode ist an ihren Grenzen dadurch charakterisiert, daß nur Töne der Tonika erlaubt sind; an allen anderen Stellen der Periode sind dagegen zusätzlich auch andere Harmonien erlaubt. Der Vordersatz V zeichnet sich dadurch aus, daß er nur auf der Tonika oder auf der 2. oder 5. Stufe mit Dominantenfunktion enden kann, die größte syntaktische Einheit darstellt, die sequenziert werden kann (ordinale Transformation s) und identisch oder mit Abwandlungen am Ende wiederholt werden kann (ordinale Transformationen $w=$ bzw. $w\approx$), wodurch zwei Vordersätze V_1 und V_2 gebildet werden.

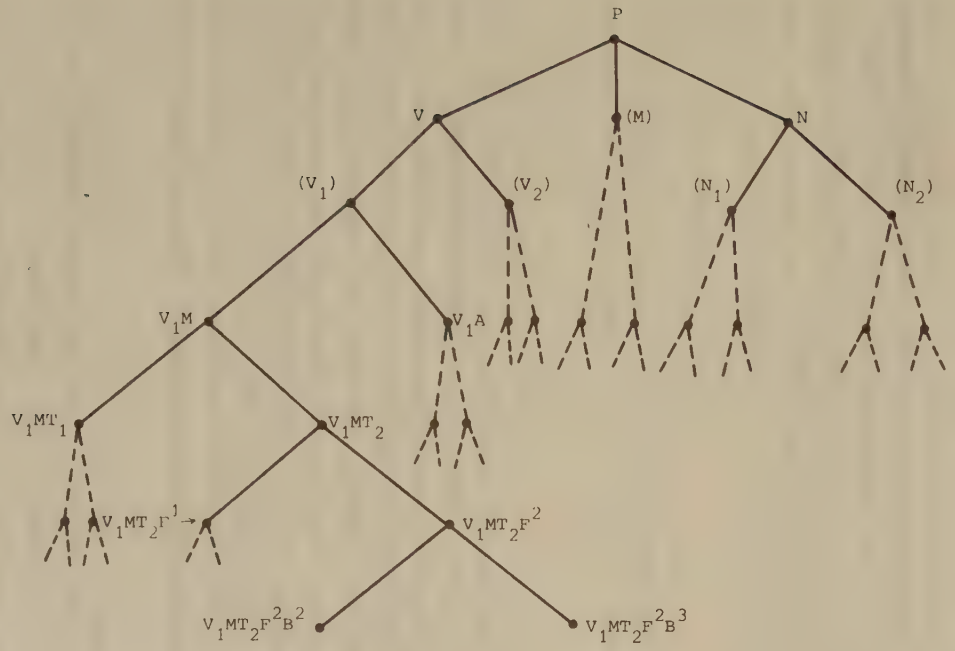
Der Mittelsatz ist die größte optionale syntaktische Einheit; an seinem Anfang können alle harmonischen Funktionen vorkommen sowie alle auf betonter Zählzeit erlaubten harmoniefremden Töne;

sein Ende ist nicht sehr charakteristisch, wird aber durch die Eigenschaften des ihm folgenden Nachsatzes eindeutig markiert. Der Nachsatz kann identisch und mit Abwandlungen wiederholt werden und kann nur auf einem Ton der Tonika enden.

Vorder-, Mittel- und Nachsatz werden in Motiv (VM, MM, NM) und Anschlußglied geteilt (VA, MA, NA). Motive und Anschlußglieder sind am deutlichsten durch die auf sie bezogenen ordinalen Transformationsoperationen zu trennen. Charakteristisch ist das Fehlen der Septime am Motivende; darüber hinaus stellt der Übergang vom Motivende zum Anschlußgliedanfang die Stelle innerhalb der Periode dar, an der die größte Wahlfreiheit herrscht.

Motive und Anschlußglieder werden alleine aufgrund rhythmischer Regelmäßigkeiten in die Takte 1 und 2 (z.B. VMT_1, VMT_2 oder NAT_1, NAT_2) geteilt: die Teilung geschieht genau vor der Wiederholung der Hauptbetonung (Betonungsrang 1). Ebenso werden aufgrund rhythmischer Regelmäßigkeiten die Takte in zwei Füße geteilt, wobei der erste den Betonungsrang 1 (Hauptbetonung) enthält (z.B. VMT_1F^1 oder MAT_2F^1) und der zweite den Rang 2 (Nebenbetonung); und die Füße werden in zwei Betonungsränge geteilt, wobei der jeweils zweite Betonungsrang den Rang 3 hat. Die Betonungsränge 1 und 2 zeichnen sich dadurch aus, daß auf ihnen nur Harmonietöne oder Vorhalte erlaubt sind, während auf dem Betonungsrang 3 darüber hinaus auch andere harmoniefremde Töne erlaubt sind (Wechselnoten, Durchgangstöne, Antizipationen u. dgl.).

Zusätzlich zu den genannten Abgrenzungskriterien liefert die Verteilung von Auftakten eine mit den vorgenommenen Teilungen konforme Konstituentenhierarchie. Dasselbe gilt auch für die noch nicht erwähnten ordinalen Transformationen auf Niveaus unterhalb der Motiv- und Anschlußgliedebene.



(Periode)

(Phrasenebene, Sätze)

(optionale Erweiterung der Phrasenebene)

(Subphrasenebene, Motive und Anschlußglieder)

(Taktebene)

(Fußebene)

(Zählzeitenebene mit den Betonungsrängen)

Abb. 1: Ausschnitt aus einem Konstituentenstrukturbaum (Bedeutung der Abkürzungen siehe Text)

1.2 Die Regeln der Syntax

1.2.1 Regeltypen

Wir werden im folgenden Abschnitt drei Typen von syntaktischen Regeln anwenden: 1. die transformationellen Ersetzungsregeln (Abk. TE-Regeln), 2. die Transformationsregeln (Abk. T-Regeln) und 3. die Ersetzungsregeln (E-Regeln).

Definition der TE-Regel: $X \rightarrow Y \left\langle \begin{matrix} t_1 \\ \vdots \\ t_i \\ \vdots \\ t_n \end{matrix} \right\rangle Z$

Dieser Regeltypus ist folgendermaßen zu lesen: die Konstituentenklasse X, die die Konstituentenklassen Y und Z dominiert, wird durch Y und Z ersetzt, wobei die Konstituentenklassen Y und Z in einer von mehreren möglichen Transformationsbeziehungen t_i zueinanderstehen können, die durch T-Regeln definiert sind. Dabei ist die Strukturbasis SB_{ti}, auf die die T-Regeln T_{ti} angewendet werden, die links von der Transformationsoperation t_i stehende Struktur, die je nach Hierarchieniveau aus nonterminalen oder aus terminalen Elementen bestehen kann. Die Strukturveränderung SV_{ti}, die nach Anwendung einer Regel T_{ti} entsteht, ist die rechts von der Transformationsoperation t_i stehende Struktur.

Der zweite Regeltyp besteht in den Transformationsregeln T_{ti}, die eine Strukturbasis SB_{ti} in eine Strukturveränderung SV_{ti} bei Anwendung der Operation t_i überführen.

Definition der T-Regeln mit einer Operation t_i :

T_{t_i} : $SB_{t_i} = \{Y\}$; $SV_{t_i} = \{t_i(Y)\}$

Dieser Regeltypus ist folgendermaßen zu lesen: die Strukturbasis SB_{ti}, auf die die Transformationsoperation t_i angewendet werden soll, besteht in der durch die Konstituentenklasse Y repräsentierten Struktur, und die Strukturveränderung ergibt sich durch Anwendung der Operation t_i auf die Strukturbasis SB_{ti}.

Der dritte Regeltypus findet nur bei der Ersetzung der in der Konstituentenhierarchie auf dem niedrigsten Niveau stehenden Konstituentenklasse durch terminale Symbole Anwendung.

Definition der E-Regel: X → y z

Die Regel besagt allgemein, daß die Konstituentenklasse X, die nonterminale Symbole enthält, durch terminale Symbole y und z ersetzt wird.

1.2.2 TE-Regeln der Syntax

Die Syntax besteht in einer sechsstelligen Relation:
$S = (V_N, V_T, R_{TE}, R_E, R_T, P)$, wobei S für die Menge aller syntaktisch zulässigen Strings steht, V_N die Menge des nonterminalen Vokabulars bezeichnet, V_T die Menge des terminalen Vokabulars darstellt und R_{TE}, R_E sowie R_T die Mengen der Regeln (TE-, E- bzw. T-Regeln) repräsentieren; P ist das nonterminale Anfangssymbol zur Ableitung eines terminalen Strings. Die Elemente der Mengen V_N, V_T, R_{TE}, R_E und R_T wollen wir hier nicht formal darstellen, da es sich um zu viele Elemente handelt, als daß wir sie hier einzeln aufzählen könnten. Sie werden aber im folgenden Text jeweils benannt.

Die Reihenfolge der Regelanwendung ist durch die Numerierung der TE-Regeln festgelegt, d.h. nachdem die TE-Regel Nr. x angewendet wurde, kann als nächste Regel nur die Regel Nr. x + 1 angewendet werden. Auf diese Weise schließen wir mögliche Rekursionen in der Anwendung einiger Regeln aus. Regeln, die wiederholt angewendet werden können, wurden nicht numeriert. Die Nummer optionaler Regeln wurde eingeklammert.

Die erste Regel ersetzt die musikalische Periode P, die in einer Liedstrophe besteht, durch die Konstituentenklassen V (Vordersätze), M (Mittelsätze) und N (Nachsätze):

$$1.\quad P \rightarrow \left\langle \begin{array}{llll} \#_V\ V & \#_V & [\#_M\ M\ \#_M] & \#_N\ N\ \#_N \\ \#_V\ (V)_N & \#_V & \#_M\ M\ \#_M & \#_N\ w= \ (V)_N\ \#_N \\ \#_V\ (V)_N & \#_V & [\#_M\ M\ \#_M] & \#_N\ w\approx \ (V)_N\ \#_N \end{array} \right\rangle$$

Hierbei bedeuten: $\#_V$, $\#_M$ und $\#_N$ die Sätze V, M und N begrenzenden terminalen Elemente, die die Funktion haben, Satzanfang und Satzende zu kennzeichnen; w = ist der Operator für die T-Regel $T_{w=}$, die eine identische Wiederholung einer Strukturbasis bewirkt; w ≈ ist der Operator für die T-Regel $T_{w\approx}$, die eine im letzten Takt modifizierte Wiederholung einer Strukturbasis bewirkt; der Index an der runden Klammer besagt, daß die Konstituentenklasse, auf die die T-Regel angewendet werden soll, auch nach der Strukturveränderung die Distributionsmerkmale der durch den Index angegebenen Konstituentenklasse aufweisen soll. Dies bedeutet z.B., daß bei der Ableitung der terminalen Elemente des Vordersatzes immer berücksichtigt werden muß, daß dieser Vordersatz später als Nachsatz auftritt, so daß z.B. das Ende des Vordersatzes so abgeleitet werden muß, daß es auch als Ende eines Nachsatzes gelten kann. Da der Vordersatz aber zusätzliche Endungsmöglichkeiten im Vergleich zum Nachsatz besitzt, darf bei der Ableitung des Endes des Vordersatzes keine

derjenigen Endungen angewendet werden, die im Nachsatz nicht erlaubt sind. $(V)_N$ können wir also lesen als Vordersatz mit den distributiven Eigenschaften eines Nachsatzes. Die eckige Klammer weist darauf hin, daß die Ersetzung durch die Konstituentenklasse M optional ist; die geschweifte Klammer weist darauf hin, daß es sich um eine Menge alternativer Ersetzungen handelt.

Als nächste Ersetzung wird der Vordersatz V optional durch zwei Vordersätze V_1 und V_2 ersetzt:

$$(2)\quad V \rightarrow \left[\begin{array}{lll} V_1 & & V_2 \;/\#_V - o \\ (V_1)_{V_2} & s & (V_1)_{V_2} \;/\#_V - M \\ (V_1)_{V_2} & w= & (V_1)_{V_2} \left\{\begin{array}{l} /\#_V - N_1 \\ /\#_V - M \end{array}\right. \\ (V_1)_{V_2} & w\approx & (V_1)_{V_2} \;/\#_V - M \end{array}\right]$$

Die Transformation s bezeichnet die in der Musiktheorie als Sequenz bezeichnete Wiederholung einer Tonfolge auf einer höheren oder niedrigeren Tonstufe. Hinter dem Schrägstrich sind die Kontexte angegeben, in denen die Transformationen auftreten können. Eine solche Kontextangabe ist formal folgendermaßen aufgebaut: hinter dem Schrägstrich folgt dasjenige terminale oder nonterminale Element, das den ersetzten Konstituentenklassen vorausgeht, dann folgt der waagerechte Strich, der stellvertretend für die ersetzte Konstituentenklasse steht, und schließlich folgt dasjenige terminale oder nonterminale Element, das den ersetzten Konstituentenklassen nachfolgt. Ein solcher Kontext ist folgendermaßen zu lesen: die durch den Kontext spezifizierte Ersetzung darf nur in diesem Kontext stattfinden, wenn der Kontext vorliegt, muß sie aber nicht stattfinden, wenn andere Alternativen noch offenstehen. Das Symbol o in einer Kontextangabe bedeutet, daß

an dieser Stelle alle Kontexte erlaubt sind. Man spricht hierbei von einem offenen Kontext. Kontexte können einseitig offen sein, wie in Regel (2), sie können aber auch beidseitig offen sein, was durch die Kontextangabe / o - o zum Ausdruck gebracht wird. Fehlt eine Kontextangabe, dann handelt es sich immer um den beidseitig offenen Kontext.

Wird durch Regel (2) die Ersetzung $(V_1)_{V_2}$ w = $(V_1)_{V_2}$ / $\not\#_V$ - N_1

vorgenommen, dann muß bei späteren Regelanwendungen dafür gesorgt werden, daß die Konstituentenklasse N_1 auch generiert wird. Dies geschieht durch Regel (3), die zwar optional ist, aber dann obligatorisch wird, wenn vorher eine Ersetzung vorgenommen wurde, die den Kontext N_1 verlangt.

(3) $N \rightarrow$

$N_1 \quad N_2 \quad / V - \not\#_N$

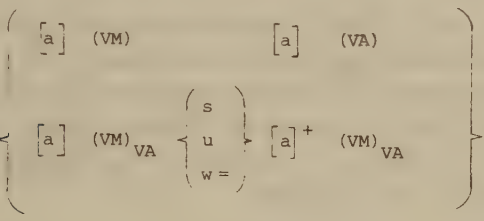
$(N_1)_{N_2} \quad w = (N_1)_{N_2}$ { $/ w = (V_1)_{V_2} - \not\#_N$; $/ V - \not\#_N$ }

$w = (V_2)_N \quad / o - \not\#_N$

$w \approx (V_2)_N \quad / M - \not\#_N$

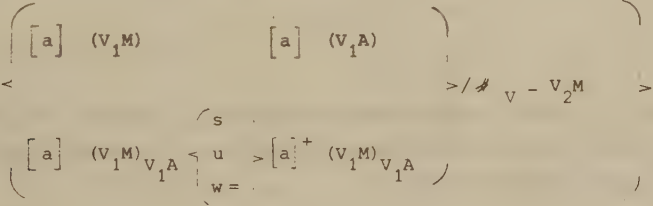
$w = (V)_N \quad / M - \not\#_N$

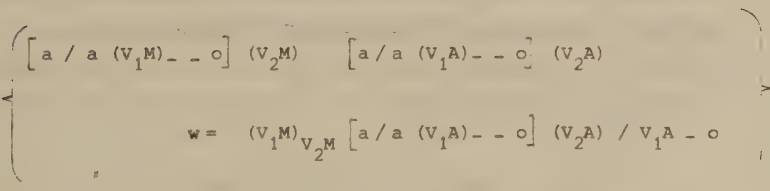
Der Vordersatz V wird in Regel 4. durch die Konstituentenklasse VM (Motiv des Vordersatzes) und durch die Konstituentenklasse VA (Anschlußglied des Vordersatzes) ersetzt:



Das in den eckigen Klammern stehende a bezeichnet den Operator für die Auftakttransformationsregel, die optional ist; das in der geschweiften inneren Klammer stehende u bezeichnet den Operator für die Transformationsregel, die eine Umkehrung der Reihenfolge der terminalen Elemente von VM bewirkt. Die Transformationsregeln werden im Abschnitt 4.2.4 dargestellt.

Die der vierten Regel entsprechenden Regeln für V_1 und V_2 sind als Regel 5. und 6. dargestellt:





a/a (V_1M)_ _ o bedeutet, daß die Auftakttransformation nur dann angewendet werden kann, wenn der vorausgehende Kontext a(V_1M) vorliegt. Das Pluszeichen an der eckigen Klammer in der zweiten Zeile der Regeln 4. und 5. besagt, daß die Auftakttrans-

formation dann obligatorisch ist, wenn das vorausgehende gleichnamige Element einen Auftakt besitzt.

Die Regeln (7)' und (8)' geben optionale Erweiterungsmöglichkeiten für die Vordersätze an, wobei die geklammerten Indizes angeben, um welchen Vordersatz es sich handelt: die leere Klammer bezeichnet V, die geklammerten Indizes 1 und 2 die Vordersätze V_1 und V_2. Regel (7)' erweitert den Vordersatz durch Einführung einer Motivsequenz:

(7)' $VM_{(\ ),(1),(2)} \rightarrow VM_{(\ ),(1),(2)} \left[a\right]^{+} s (VM_{(\ ),(1),(2)})^{VM}_{(\ ),(1),(2)}$

$/ \not\# _{V} - VA_{(\ ),(1),(2)}$

Regel (8)' erweitert die Vordersätze durch Anschlußgliedwiederholung:

(8)' $VA_{(\ ),(1),(2)} \rightarrow VA_{(\ ),(1),(2)} \left[a\right]^{+} w = (VA_{(\ ),(1),(2)})^{VA}_{(\ ),(1),(2)}$

$/ VM_{(\ ),(1),(2)} - o$

Der Strich an der Regelnummer bedeutet, daß alle anderen optionalen Erweiterungen des Vordersatzes nicht vorgenommen werden dürfen (vgl. Regel (17)).

Nun wird der Nachsatz N durch das Motiv des Nachsatzes NM und durch das Anschlußglied des Nachsatzes NA ersetzt:

9. $N \rightarrow \left\langle \begin{array}{ll} \left[a\right] (NM) & \left[a\right] (NA) \\ \left[a\right] (NM)_{NA} \left\{ \begin{matrix} s \\ u \end{matrix} \right\} & \left[a\right]^{+} (NM)_{NA} \\ \left[a\right]^{+} w = (VM)_{NM} & \left[a\right] (NA) \\ \left[a\right]^{+} w = (w = (V_1A)_{V_2A})_{NM} & \left[a\right] (NA) \\ \left[a\right]^{+} w = (VA)_{NM} & \left[a\right] (NA) \\ \multicolumn{2}{c}{\left[a\right]^{+} w = (V_2A)_{NA} \ / \ MA - \not\# _{N}} \end{array} \right\rangle$

Die der 9. Regel entsprechenden Regeln für N_1 und N_2 sind die folgenden Regeln 10. und 11.:

10. $N_1 \rightarrow$

$[a]\ (N_1M)$ $\quad$ $[a]\ (N_1A)$

$[a]\ (N_1M)_{N_1A} \left\{ \begin{matrix} s \\ u \\ w= \end{matrix} \right\}$ $\quad$ $[a]^+ (N_1M)_{N_1A}$

$[a]^+\ w = (VM)_{N_1M}$ $\quad$ $[a]\ (N_1A)$ $\quad$ $\Big\} /o - N_2M$

$[a]^+\ w = (w = (V_1A)_{V_2A})_{N_1M}$ $\quad$ $[a]\ (N_1A)$

$[a]^+\ w = (VA)_{N_1M}$ $\quad$ $[a]\ (N_1A)$

$[a]^+\ w = (w = (V_1M)_{V_2M})_{N_1M}$ $\quad$ $[a]\ (N_1A)/o_w = (w=(w=(V_1M)_{V_2M})_{N_1M})_{N_2M}$

11. $N_2 \rightarrow$

$[a/a(N_1M) - - o](N_2M)$ $\quad$ $[a/a(N_1A) - - o]\ (N_2A)$ $\quad$ $\Big\} /o - \not\emptyset_N$

$[a]^+\ w = (N_1M)_{N_2M}$ $\quad$ $[a/a(N_1A) - - o]\ (N_2A)$

$[a]^+\ w = (w=(w=(V_1M)_{V_2M})_{N_1M})_{N_2M}$ $\quad$ $[a]\ (N_2A)/w=(w=(V_1M)_{V_2M})_{N_1M}{}^{N_1A - \not\emptyset_N}$

Es folgen nun die den Regeln (7)' und (8)' entsprechenden Regeln für die optionale Erweiterung der Nachsätze. Regel (12)' erweitert die Nachsätze durch Motivwiederholung oder durch Motivsequenz:

(12)' $NM_{(\ ),(1),(2)} \rightarrow NM_{(\ ),(1),(2)}\ [a]^+ \left\{ \begin{matrix} w= \\ s \end{matrix} \right\} (NM_{(\ ),(1),(2)})_{NM_{(\ ),(1),(2)}}$

$/o - NA_{(\ ),(1),(2)}$

Regel (13)' erweitert die Nachsätze N und N_1 sowie N_2 durch Anschlußgliedsequenz:

(13)' $NA_{(),(2)} \rightarrow$ ($NA_{(),(2)}$ [a]$^+$ s($NA_{(),(2)}$)$NA_{(),(2)}$ /$NM_{(),(2)}$ - # N ; N_1A [a]$^+$ s(N_1A)N_1A /N_1M - N_2M)

Hier bedeutet der Strich an der Regelnummer analog zu den Regeln (7)' und (8)', daß alle weiteren optionalen Erweiterungen der Nachsätze dann nicht angewendet werden dürfen, wenn diese Regeln angewendet wurden.

Regel 14. ersetzt den Mittelsatz M durch das Motiv des Mittelsatzes MM und durch das Anschlußglied des Mittelsatzes MA:

14. M → ([a] (MM) [a/a(MM) - o] (MA) ; [a]$^+$ (MM)$_{MA}$ < s / w= / u > [a]$^+$ (MM)$_{MA}$; [a] (MA) / V_2A - NM)

Damit ist die Ersetzung aller Sätze durch Motiv und Anschlußglied abgeschlossen. Nun folgt die Ersetzung der Motive und Anschlußglieder in die Motiv- bzw. Anschlußgliedhälften, die als Takte 1 und 2 der Motive bzw. Anschlußglieder bezeichnet werden.

15. $VM_{(),(1),(2)} \rightarrow$ ($VMT_{1(),(1),(2)}$ [a/a($VMT_{1(),(1),(2)}$)-a($VAT_{1(),(1),(2)}$)] $VMT_{2(),(1),(2)}$; ($VMT_{1(),(1),(2)}$)$VMT_{2(),(1),(2)}$ s / u [a]$^+$ ($VMT_{1(),(1),(2)}$)$VMT_{2(),(1),(2)}$; ($VMT_{1(),(1),(2)}$)$VMT_{2(),(1),(2)}$ [a]$^+$ w=($VMT_{1(),(1),(2)}$)$VMT_{2(),(1),(2)}$; / # $_V$ - $VAT_{1(),(1),(2)}$)

16. $VA_{(),(1),(2)} \rightarrow \left\langle \begin{array}{l} VAT_{1(),(1),(2)} \left[a/a(VAT_{1(),(1),(2)}) - o\right] \; VAT_{2(),(1),(2)} \\ VAT_{1(),(1),(2)})VAT_{2(),(1),(2)} \begin{Bmatrix} s \\ u \end{Bmatrix} [a]^{+} (VAT_{1(),(1),(2)})VAT_{2(),(1),(2)} \end{array} \right\rangle$

Regel (17)' ist eine optionale Erweiterung der Motive in den Vordersätzen, die nur dann stattfinden kann, wenn weder Regel (7)' noch Regel (8)' angewendet wurden:

(17)' $VMT_{2(),(1),(2)} \rightarrow VMT_{2(),(1),(2)} [a]^{+} w = (VMT_{2(),(1),(2)})VMT_{2(),(1),(2)}$

$/ \; VMT_{2(),(1),(2)} - VAT_{1(),(1),(2)}$

18. $MM \rightarrow \left\langle \begin{array}{l} MMT_1 \quad \left[a/a(MMT_1) - a(MAT_1)\right] \quad (MMT_2) \\ (MMT_1)_{MMT_2} \quad \begin{Bmatrix} s \\ u \end{Bmatrix} \quad [a]^{+} \quad (MMT_1)_{MMT_2} \end{array} \right\rangle$

19. $MA \rightarrow \left\langle \begin{array}{l} MAT_1 \quad \left[a/a(MAT_1) - o\right] \quad (MAT_2) \\ (MAT_1)_{MAT_2} \quad \begin{Bmatrix} s \\ u \end{Bmatrix} \quad [a]^{+} \quad (MAT_1)_{MAT_2} \end{array} \right\rangle$

20. $NM_{(),(1),(2)} \rightarrow \left\langle \begin{array}{l} NMT_{1(),(1),(2)} \qquad [a] \; (NMT_{2(),(1),(2)}) \\ (NMT_{1(),(1),(2)})NMT_{2(),(1),(2)} \begin{Bmatrix} s \\ u \end{Bmatrix} [a]^{+} (NMT_{1(),(1),(2)})NMT_{2(),(1),(2)} \\ (NMT_{1(),(1),(2)})NMT_{2(),(1),(2)} [a]^{+} w = (NMT_{1(),(1),(2)})NMT_{2(),(1),(2)} \\ / \; o - NAT_1 \end{array} \right\rangle$

Regel (21) ist eine optionale Erweiterung des Motivs in dem Nachsatz N, die nur dann stattfinden kann, wenn weder Regel (12)' noch Regel (13)' angewendet wurden:

$$(21)' \quad NMT_2 \rightarrow NMT_2 \; [a]^{+} \; w = (NMT_2)_{NMT_2} \quad / \; o \; - \; NAT_1$$

$$22. \quad NA_{(),(1),(2)} \rightarrow \left\langle \begin{array}{lll} NAT_{1(),(2)} & [a] & NAT_{2(),(2)} \quad / \; o \; - \not{N} \\ N_1AT_1 & [a] & N_1AT_2 \quad / \; o \; - \; N_2AT_1 \end{array} \right\rangle$$

Die Ersetzung aller Motive und Anschlußglieder durch Takte ist nun abgeschlossen. Es folgt jetzt die Festlegung der Betonungen innerhalb eines Taktes, die von dem gewählten Metrum abhängt, sowie die Ersetzungsregeln für Füße durch terminale Elemente. Gleichzeitig mit der Ersetzung der Takte durch den 1. und 2. Fuß werden die Betonungsränge eines Taktes generiert. Dies führen wir zunächst für das Beispiel der Ersetzung des 1. Taktes des Vordersatzmotivs beim 4/4-Takt aus und verallgemeinern dann auf die weiteren Takte und andere Metren.

$$VMT_{1(),(1),(2)} \rightarrow \left\langle \begin{array}{l} VMT_1F^1{}_{(),(1),(2)} \quad VMT_1F^2{}_{(),(1),(2)} \\ VMT_1F^1{}_{(),(1),(2)} \left\{ \begin{matrix} u \\ w= \\ s \end{matrix} \right\} (VMT_1F^1{}_{(),(1),(2)})_{VMT_1F^2{}_{(),(1),(2)}} \end{array} \right\rangle$$

Der hochgestellte Index gibt den ersten Betonungsrang eines Fußes an. Je niedriger ein Betonungsrang, desto stärker ist die Betonung der ersten Zählzeit innerhalb eines Fußes. Die Betonungsränge ergeben sich aus der Anwendung des "left priority rule" (CHOMSKY & HALLE, 1968): jeder Takt enthält den Betonungsrang 1, und beim Ersetzen eines Taktes durch zwei Füße wird der auf glei-

cher Hierarchiestufe am weitesten links stehende Rang, also der des 1. Fußes, beibehalten, während auf gleicher Hierarchiestufe rechts davon stehende Füße (im Fall des 4/4-Taktes der 2. Fuß) einen um eine Einheit höheren Rang erhalten. Werden die Füße schließlich durch terminale Elemente ersetzt, wird noch einmal das "left priority rule" angewendet. Beim 4/4-Takt haben wir vier Zählzeiten. Jede dieser Zählzeiten erhält zunächst den Betonungsrang 1. Bei Ableitung der Füße aus dem Takt und bei Ableitung der terminalen Elemente (ein Ton pro Zählzeit) wird jeweils das "left priority rule" angewendet. Die folgende Klammerdarstellung beschreibt die Generierung der Betonungsränge beim 4/4-Takt formal:

$$((1)(1))_{F^1}\,((1)(1))_{F^2})_T \rightarrow (({1\,2})_{F^1}\,({1\,2})_{F^2})_T \rightarrow (1\,3\,2\,3)_T$$

Analog verfahren wir beim 3/4-Takt, wobei wir ebenfalls von der Ableitung zweier Füße ausgehen, weil die dritte Zählzeit als nebenbetont betrachtet werden kann und außerdem die Distribution der Harmonien und der harmoniefremden Töne für die dritte Zählzeit im 3/4-Takt der der dritten Zählzeit im 4/4-Takt entspricht:

$$((1)(1))_{F^1}\,(1)_{F^2})_T \rightarrow (({1\,2})_{F^1}\,(1)_{F^2})_T \rightarrow (1\,3\,2)_T$$

Der 6/8-Takt wird dann wie folgt mit Betonungsrängen versehen:

$$(((1)(1))(1))_{F^1}\,(((1)(1))\,(1))_{F^2})_T \rightarrow$$

$$((1\,2)\,(1))_{F^1}\,((1\,2)\,(1))_{F^2})_T \rightarrow (({1\,3\,2})_{F^1}\,({1\,3\,2})_{F^2})_T \rightarrow (1\,4\,3\,2\,4\,3)_T$$

Für alle Zweiertakte können wir die Betonungsränge verwenden, die für den 4/4-Takt abgeleitet wurden. Die Viererbetonungsränge ergeben sich beim 6/8-Takt durch die dreimalige Anwendung des "left priority rule".

Für die Ersetzung aller Takte durch Füße gilt die folgende Regel:

23. $- T \rightarrow \left\{ \begin{array}{ll} - F^1 & - F^2 \\ - F^1 \left\{ \begin{array}{l} u \\ w= \\ s \end{array} \right. & - F^2 \end{array} \right\}$

wobei der Strich vor dem T und dem F als Stellvertreter für die Buchstaben steht, die die Konstituenten näher bezeichnen.

1.2.3 E - Regeln der Syntax

Die nun folgenden Regeln sind E-Regeln für die Ersetzung der Füße durch terminale Elemente. Für die terminalen Elemente verwenden wir im Unterschied zur Standardsyntaxtheorie ebenfalls Großbuchstaben, weil die terminalen Elemente neben der Tonstufe innerhalb einer diatonischen Tonleiter auch die harmonische Funktion des Tones zum Ausdruck bringen und es in der Musiktheorie üblich ist, Harmonien durch Großbuchstaben zu symbolisieren. Die harmonischen Funktionen sind durch die Tonika (T), die Subdominante (S) und die Dominante (D) symbolisiert; die jeweiligen Tonstufen durch Indizes. Z.B. bedeutet T_1, daß es sich um das terminale Element handelt, das die erste Tonstufe der diatonischen Tonleiter bildet und das die harmonische Funktion der Tonika besitzen soll; bei S_1 dagegen besitzt derselbe Ton die harmonische Funktion der Subdominante. Zu den terminalen Elementen gehören auch solche, die keine harmonische Funktion besitzen, also harmoniefremde Töne darstellen.

Harmoniefremde Töne sind Durchgangstöne, Wechseltöne, Vorhalte, Antizipationen und Einführungen. Die folgenden Regeln definieren diese harmoniefremden Töne formal.

a) Durchgangstöne

Durchgangstöne sind terminale Elemente auf unbetonter Zählzeit (Betonungsränge 3 oder größer), denen ein in der diatonischen Tonleiter unmittelbar benachbarter höherer (tieferer) Harmonie-Ton auf betonter Zählzeit (Betonungsränge 1 und 2) vorausgeht und ein in der diatonischen Tonleiter unmittelbar benachbarter tieferer (höherer) Harmonie-Ton auf betonter Zählzeit folgt:

$$Dg_i : \left\{ \begin{array}{cccc} (1) & & & \\ (2) & (3) & & \\ H^x_{i-1} & Dg_i & / \circ - & H^y_{i+1} \\ H^x_{i+1} & Dg_i & / \circ - & H^y_{i-1} \end{array} \right\}$$

für x und y $\in \{T, S, D\}$ und i $\in \{1, 2, \ldots, 6, 7\}$. Dg_i bezeichnet den Durchgangston auf der Tonstufe i , H^x und H^y sind beliebige Harmonietöne, deren harmonische Funktion durch die hochgestellten Indizes x und y symbolisiert werden. Die eingeklammerten Zahlen geben die Betonungsränge an.

b) Wechseltöne

Wechseltöne W_i sind terminale Elemente der Tonstufe i auf unbetonter Zählzeit, denen ein in der diatonischen Tonleiter unmittelbar benachbarter höherer (tieferer) Harmonie-Ton auf betonter Zählzeit vorausgeht und derselbe Harmonie-Ton nachfolgt:

$$W_i : \left\{ \begin{array}{cccc} (1) & (3) & & \\ H^x_{i-1} & W_i & / \circ - & H^x_{i-1} \\ H^x_{i+1} & W_i & / \circ - & H^x_{i+1} \end{array} \right\}$$

c) Vorhalte

Vorhalte V_i sind terminale Elemente der Tonstufe i auf betonter Zählzeit, denen ein in der diatonischen Tonleiter unmittelbar benachbarter tieferer Harmonieton auf unbetonter Zählzeit folgt:

$$V_i : \overset{(1)}{V_i} \quad \overset{(3)}{H^x_{i-1}} \quad / o - H^x_j$$

für i und j $\in$ {1, 2, ..., 6, 7}

d) Antizipationen

Antizipationen A_i sind terminale Elemente der Tonstufe i auf unbetonter Zählzeit, denen auf betonter Zählzeit ein Harmonieton derselben Tonstufe i folgt:

$$A_i : \overset{(2)}{H^x_j} \quad \overset{(3)}{A_i} \quad / o - H^y_i \quad \text{mit} \quad x \neq y$$

e) Einführungen

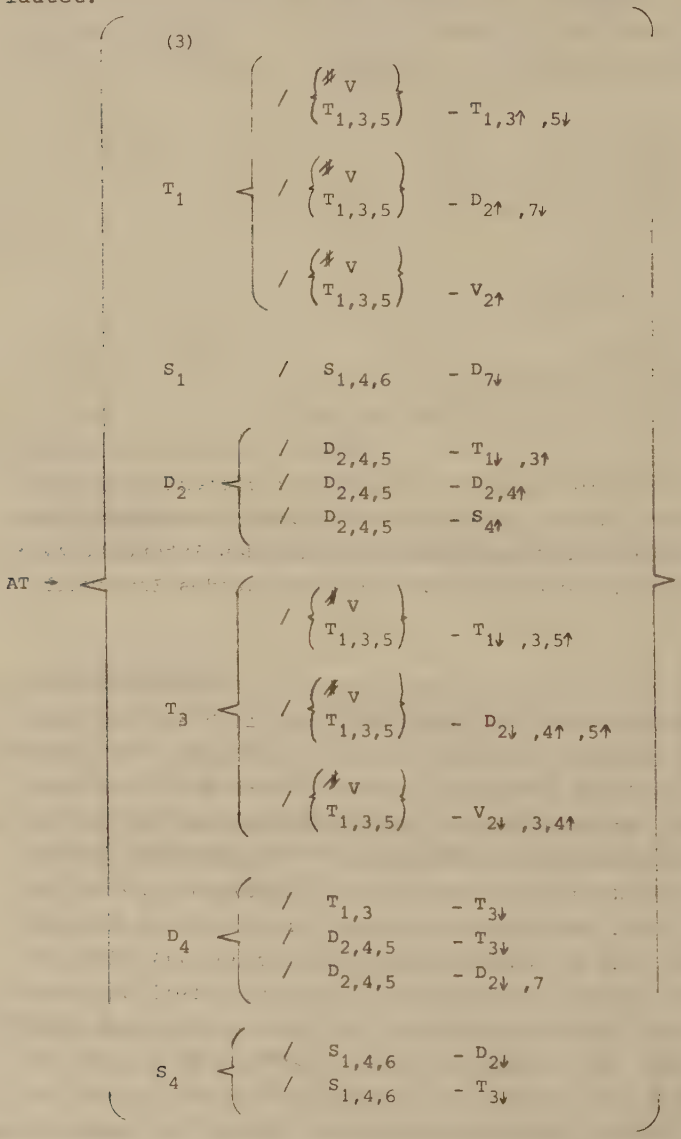
Einführungen E_i sind terminale Elemente der Tonstufe i auf unbetonter Zählzeit, denen auf betonter Zählzeit ein Harmonieton vorausgeht, dessen harmonische Funktion auf die Tonstufe i nicht anwendbar ist, und denen auf betonter Zählzeit ebenfalls ein Harmonieton folgt, dessen harmonische Funktion auf die Tonstufe i ebenfalls nicht anwendbar ist:

$$E_i : \left\{ \begin{array}{lll} (3) & (2) & (1) \\ E_2 & / H_j^x - D_{7\downarrow} & \\ E_{2,3} & / H_j^x - T_{1\downarrow} & \\ E_4 & \left\{ \begin{array}{l} / H_j^x - T_{3\downarrow} \\ / H_j^x - D_{2\downarrow} \end{array} \right. & \\ E_5 & / H_j^x - T_{3\downarrow} & \\ E_6 & / H_j^x - T_{5\downarrow} & \end{array} \right\}$$

wobei $i \neq j$ ist. Hier haben wir gleich nur diejenigen Einführungen aufgezählt, die in unserem Liedercorpus vorgekommen sind. Der Pfeil in der Kontextangabe bedeutet, daß das Intervall zwischen E_i und dem nachfolgenden Ton ein absteigendes Intervall sein muß.

Durch die am Schluß noch zu formalisierende Auftakttransformation wird eine Konstituentenklasse generiert, die sich auf dem gleichen Niveau in der Konstituentenhierarchie befindet, wie die Füße. Weil der erste Auftakt derjenige zum Vordersatz sein kann, muß die den Auftakt repräsentierende Konstituentenklasse AT als erste durch terminale Elemente ersetzt werden. Da diese Regel an mehreren Stellen der Konstituentenhierarchie angewendet werden muß, haben wir ihr keine Regelnummer gegeben. Die Pfeile im Kontext bedeuten wieder, daß es sich um ein aufsteigendes (↑) oder absteigendes (↓) Intervall handeln muß; wird kein Pfeil angegeben, dann kann es sich sowohl um ein auf- als auch um ein absteigendes Intervall handeln. Da Intervalle größer als eine Sext nicht erlaubt sind, handelt es sich bei Indexgleichheit von Auftakt und nachfolgendem Ton um identische Tonstufen (Oktave ist ausgeschlossen).

Die Ersetzungsregel für die Konstituentenklasse der Auftakte AT lautet:



Die Fortsetzung der Ersetzungsregel für die Auftakte befindet sich auf der nächsten Seite.

(3)

AT →

T_5 :
- / $\left\{ \begin{matrix} \not{*}\, V \\ T_{1,3,5} \end{matrix} \right\}$ - $T_{1,3\downarrow\ ,5}$
- / $T_{1,3,5}$ - $D_{4\downarrow\ ,5}$
- / $T_{1,3,5}$ - $S_{6\uparrow}$
- / $\left\{ \begin{matrix} \not{*}\, V \\ T_{1,3,5} \end{matrix} \right\}$ - $V_{3\downarrow\ ,6\uparrow}$

D_5 :
- / $D_{2,4,5}$ - $T_{1,3\downarrow\ ,5}$
- / $D_{2,4,5}$ - D_5
- / $D_{2,4,5}$ - $S_{6\uparrow}$

S_6 : / $S_{1,4,6}$ - $T_{5\downarrow}$

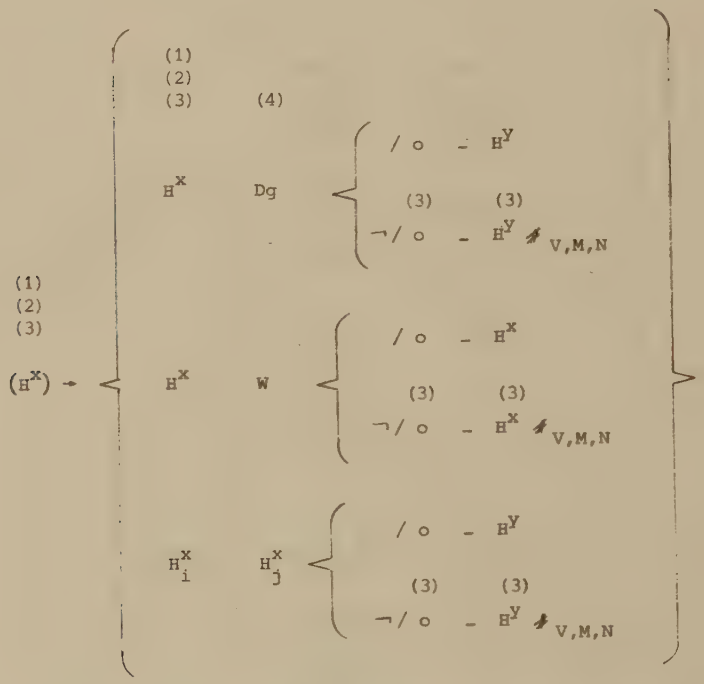
D_7 : / $D_{2,4,5}$ - D_2

E_2 : / H_j^x - $T_{1\downarrow}$ für $j \neq 2$

E_4 : / H_j^x - $T_{3\downarrow}$ für $j \neq 4$

E_6 :
- / H_j^x - $T_{5\downarrow}$
- / $\not{*}\, V$ - $T_{5\downarrow}$ für $j \neq 6$
- / H_j^x - $D_{5\downarrow}$

Ebenso wie die Ersetzungsregel für den Auftakt ist die Regel für die Besetzung der Betonungsränge von 4 eine Regel, die nicht an eine festgelegte Reihenfolge der Regelanwendungen gebunden sein kann, weil sie für jeden Fuß der ersten drei Takte eines Satzes und für den ersten Fuß des vierten Taktes eines Satzes angewendet werden kann. Deshalb erhält diese optionale Regel keine Nummer, das zu ersetzende terminale Element wird aber in runde Klammern gesetzt, um zu kennzeichnen, daß es sich hierbei um eine optionale Regel handelt.



Das Negationszeichen vor der Kontextangabe bedeutet, daß die Ersetzung in diesem Kontext nicht vorgenommen werden kann. Hier wird ausgeschlossen, daß im zweiten Fuß des vierten Taktes eines Satzes Betonungsränge von 4 auftreten. Die eingeklammerten Zahlen über den terminalen Elementen geben die Betonungsränge an, für

die eine Ersetzung möglich ist bzw. innerhalb der Kontextangabe geben sie den Betonungsrangkontext an, in dem eine Ersetzung nicht erlaubt ist.

In den nun gleich folgenden Ersetzungsregeln für die Ersetzung der Füße durch je zwei terminale Elemente sind umfangreiche Kontextangaben notwendig, weil die Aufeinanderfolge einiger Tonstufen ausgeschlossen werden muß. Es sind z.B. keine Intervalle erlaubt, die größer als eine Sext sind. Die Distributionsanalyse der Tonfolgen hat ergeben, daß auch einige Tonstufenfolgen, die ein Intervall kleiner als eine Sext bilden, nicht im Corpus vorkommen; z.B. darf der siebten Tonstufe (D_7) nicht die Terz der Tonika (T_3) folgen. Um die Kontextangaben nicht zu umfangreich und unübersichtlich werden zu lassen, haben wir als eine Art Metakontext Regeln formuliert, die die erlaubten Kontexte für alle Tonstufen der diatonischen Tonleiter angeben. In den Ersetzungsregeln wird jetzt in den meisten Fällen auf eine Indizierung der Harmonien verzichtet, wenn die Metakontexte, die für alle Füße gelten, als Kontextangaben ausreichen. Nur in den Fällen, in denen die erlaubten Kontexte noch stärker eingeschränkt sind, als es die Metakontexte angeben, werden die Harmonien indiziert und die eingeschränkteren Kontexte angegeben.

Beispiel: Die folgende Ersetzungsregel sei angegeben

$$VMT_2\ F^1 \rightarrow \overset{(1)}{T}\ \overset{(3)}{T}\ /\ D\ _\ T$$

Dies heißt ohne Metaregeln für Kontexte, daß der erste Fuß des zweiten Taktes im Vordersatzmotiv auf Betonungsrang (1) sowie auf Betonungsrang (3) irgendeinen Tonikaton erhalten kann, wenn irgendein Ton der Dominante vorausgeht und irgendein Tonikaton nachfolgt. Mit Metaregeln für Kontexte heißt dieselbe Formel,daß ein solcher Tonikaton auf dem ersten Betonungsrang stehen kann, der im Kontext des vorausgehenden Dominantentons erlaubt ist,

und daß nur ein solcher Tonikaton auf dem Betonungsrang (3) stehen kann, der im Kontext des nachfolgenden Tonikatons erlaubt ist. Außerdem wird durch die Metakontexte auch die Folge der beiden Ersetzungselemente (im Beispiel die beiden Tonikatöne) geregelt, das heißt, daß hier nur solche Tonikatöne auf die Betonungsränge (1) und (3) gesetzt werden dürfen, die miteinander als Kontext für den jeweiligen anderen Ton vorkommen dürfen.

Ist nun der vorausgehende Kontext die Septime der diatonischen Tonleiter mit der harmonischen Funktion der Dominante (D_7), dann ist als nachfolgender Tonikaton auf Betonungsrang (1) nicht die dritte Tonstufe der Tonika (T_3) erlaubt, auf Betonungsrang (3) aber kann jeder beliebige Tonikaton folgen, weil die Tonikatöne alle miteinander frei kombinierbar sind.

Die Metaregeln zur Einschränkung der Tonfolgekombinationen sehen allgemein formuliert folgendermaßen aus:

$$H_i^x : \left\langle \begin{array}{l} / - H_i^x \\ / - H_j^y \end{array} \right\rangle$$

Diese Regel muß wie folgt gelesen werden: Die Harmonie H^x mit der Funktion x und der Tonstufe i darf nur in den folgenden Kontexten erscheinen: entweder muß H_i^x die gleiche Harmonie H_i^x folgen, oder es darf die Harmonie H_j^y mit der Funktion y und der Tonstufe j folgen, alle anderen Kontexte für nachfolgende Töne sind nicht erlaubt. Für vorausgehende Kontexte gilt die folgende allgemeine Schreibweise:

$$H_i^x : \left\langle \begin{array}{l} / H_i^x - \\ / H_j^y - \end{array} \right\rangle$$

Bei der Anwendung der Ersetzungsregeln zur Erzeugung terminaler Elemente ist also für jeden Ton, gleich an welcher Stelle in der musikalischen Periode, zu prüfen, ob der vorausgehende und der nachfolgende Ton die Anwendung einer Ersetzungsregel gestattet, es sei denn, die in der Ersetzungsregel angegebenen Indizes schränken die Tonfolgekombinationen über die nach den Metaregeln erlaubten Kontexte hinaus ein.

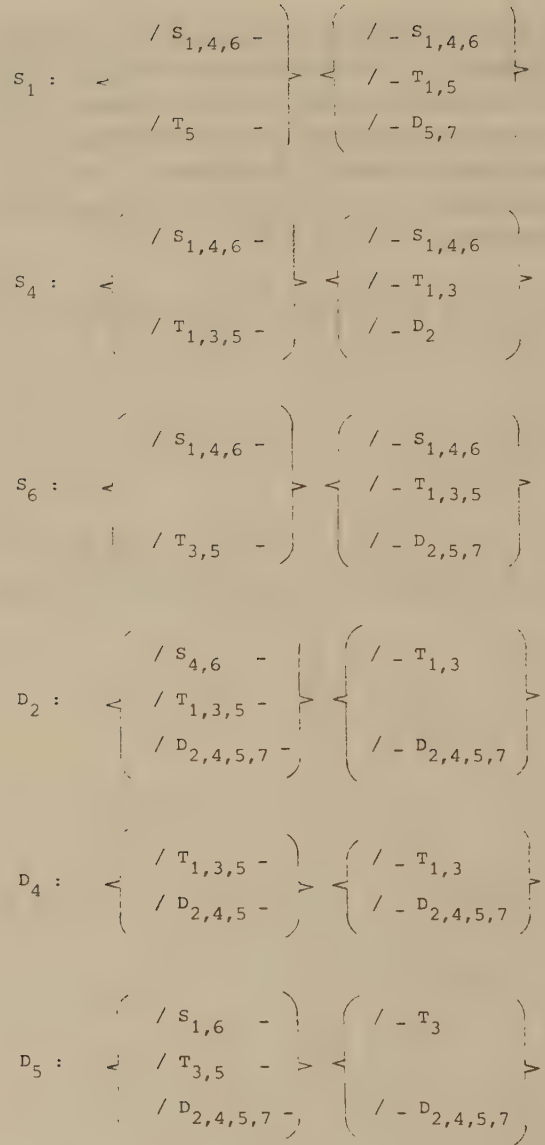
Nun folgen die Metaregeln für jede Tonstufe der diatonischen Tonleiter:

$$T_1 : \left\langle \begin{matrix} / S_{1,4,6} - \\ / T_{1,3,5} - \\ / D_{2,4,7} - \end{matrix} \right\rangle \left\langle \begin{matrix} / - S_4 \\ / - T_{1,3,5} \\ / - D_{2,4,7} \end{matrix} \right\rangle$$

Die vorausgehenden und nachfolgenden Kontexte sind frei miteinander kombinierbar.

$$T_3 : \left\langle \begin{matrix} / S_{4,6} - \\ / T_{1,3,5} - \\ / D_{2,4,5} - \end{matrix} \right\rangle \left\langle \begin{matrix} / - S_{4,6} \\ / - T_{1,3,5} \\ / - D_{2,4,7} \end{matrix} \right\rangle$$

$$T_5 : \left\langle \begin{matrix} / S_{1,6} - \\ / T_{1,3,5} - \\ / D_{2,4,7} - \end{matrix} \right\rangle \left\langle \begin{matrix} / - S_{1,4,6} \\ / - T_{1,3,5} \\ / - D_{2,4,5} \end{matrix} \right\rangle$$


S1 :
/ S1,4,6 -
/ T5 -
/ - S1,4,6
/ - T1,5
/ - D5,7
S4 :
/ S1,4,6 -
/ T1,3,5 -
/ - S1,4,6
/ - T1,3
/ - D2
S6 :
/ S1,4,6 -
/ T3,5 -
/ - S1,4,6
/ - T1,3,5
/ - D2,5,7
D2 :
/ S4,6 -
/ T1,3,5 -
/ D2,4,5,7 -
/ - T1,3
/ - D2,4,5,7
D4 :
/ T1,3,5 -
/ D2,4,5 -
/ - T1,3
/ - D2,4,5,7
D5 :
/ S1,6 -
/ T3,5 -
/ D2,4,5,7 -
/ - T3
/ - D2,4,5,7

$$D_7 : \left\langle \begin{Bmatrix} / S_{1,6} \\ / T_{1,3} \\ / D_{2,4,5,7} \end{Bmatrix} - \right\rangle \left\langle \begin{Bmatrix} / - T_{1,5} \\ \\ / - D_{2,5,7} \end{Bmatrix} \right\rangle$$

Diese Metaregeln formalisieren das "Gesetz der Nähe", wonach einer bestimmten Tonstufe nur ein solcher Ton folgen darf, der innerhalb der für ihn festgelegten harmonischen Funktion dem vorausgehenden Ton möglichst nahe liegt. Diese Regeln konnten SUNDBERG & LINDBLOM (1976) nicht formalisieren, obwohl sie für die von ihnen untersuchten schwedischen Volkslieder ebenfalls gelten. Gleichzeitig ist hier auch die Kadenzregel enthalten, wonach einem Ton mit dominantischer Funktion kein Ton mit subdominantischer Funktion folgen darf. Diese Regel gilt zwar nicht generell für alle musikalischen Formen und Stile, wird aber in allen Liedern unseres Corpus eingehalten.

Als nächstes folgen die Ersetzungsregeln für die einzelnen Füße durch terminale Elemente. Wenn keine Indizes für die harmonischen Funktionen angegeben sind, dann gelten immer die folgenden Indizes, die alternativ einsetzbare Tonstufen bezeichnen:

a) für die Tonika: $T_{1,3,5}$

b) für die Subdominante: $S_{1,4,6}$

c) für die Dominante: $D_{2,4,5,7}$

Indizes werden nur dann angegeben, wenn nicht alle vorstehend aufgeführten Indizes an der betreffenden Stelle erlaubt sind. Die Kombinierbarkeit aller Indizes für terminale Elemente und Kontexte wird durch die oben beschriebenen Metaregeln zur Einschränkung der Tonfolgekombinationen geregelt. Ein Punkt neben einem Symbol für die harmonische Funktion (z.B. T•) bedeutet, daß eine optionale Punktierung des Notenwertes vorgenommen werden kann.

Ein Bogen unter zwei Symbolen für die harmonische Funktion bedeutet, daß an dieser Stelle eine optionale Bindung zweier Notenwerte eingeführt werden kann. Die Transformationsregeln hierfür werden zusammen mit den anderen Transformationsregeln am Ende dieses Kapitels unter Punkt 1.2.4 angegeben.

Die folgende allgemeine Regel bei der Interpretation der Kontextangaben muß beachtet werden. In den vorausgehenden Kontexten steht meistens nur eine Harmonie. Da aber auch harmoniefremde Töne vorausgehen können und wir dies nicht immer von neuem angeben wollten, gilt die folgende Vereinbarung: Wenn eine E-Regel angewendet wird, die keinen offenen, vorausgehenden Kontext enthält, und links von dem durch diese Regelanwendung zu ersetzenden Element steht ein harmoniefremder Ton, dann besagt die Kontextangabe $/H^x$- nicht, daß dieser Ersetzung nur ein Harmonieton mit der Funktion x vorausgehen kann, sondern daß entweder ein Harmonieton H mit der Funktion x oder ein harmoniefremder Ton vorausgehen kann, dem dann aber die Harmonie H^x vorausgehen muß.

Beispiel: Durch Regel 24 seien die beiden Elemente T_1 und W_2 generiert worden; nun soll die Regel 25, Zeile 1 angewendet werden, um eine Ersetzung von VMT_1F^2 durch T_1 und T_5 vorzunehmen; die Kontextangabe dieser Regel aber lautet: / T - o , das heißt nun, daß entweder dem T_1 die Tonika vorausgehen muß, was nicht der Fall ist, da W_2 vorausgeht, oder daß ein harmoniefremder Ton vorausgehen muß, dem selbst wieder die Tonika vorausgeht. Letzteres ist der Fall, also kann die Ersetzung durch Regel 25, Zeile 1, vorgenommen werden.

24. VMT_1F^1 (), (1), (2)

(1)	(3)	
T•	T	/ o - o
$D^{\bullet}_{2,7}$	$D_{2,7}$	/ T - D
$D^{\bullet}_{2}$	T_1	/ $T_{1,3}$ - D
$V_{2,\ 4,6}$	H^X	/ T - H^X
$H^{X}\bullet$	W	/ o - H^X
$H^{X}\bullet$	Dg	/ o - o

25. VMT_1F^2 (), (1), (2)

(2)	(3)	
T•	T	/ T - o
$D^{\bullet}_{2,4,5}$	$D_{2,4,5}$	/ $\begin{Bmatrix}T\\D\end{Bmatrix}$ - $\begin{Bmatrix}D\\T\end{Bmatrix}$
S•	S	/ T - T
T•	D	/ T - D
D•	T	/ D - $\begin{Bmatrix}T\\D\end{Bmatrix}$
S•	T	/ T - $\begin{Bmatrix}T\\D\end{Bmatrix}$
$V_{2,3,4,5,6}$	H^X	/ o - H^X
$H^{X}\bullet$	Dg	/ o - o
$H^{X}\bullet$	E	/ o - o
$H^{X}\bullet$	W	/ o - o

26. VMT_2F^1 (), (1), (2) →

(1)	(3)	
T•	T	/ T – o / D – T / S – {S, D}
D•	D	/ T – {T, D}
S•	S	/ T – o
S_6	$T_{1\uparrow}$	/ T – ♩ AT
$V_{2,3,4,5,6}$	H^X	/ o – H^X
H^X•	W	/ o – H^X
H^X•	Dg	/ o – o

27. VMT_2F^2 (), (1), (2) →

(2)	(3)	
T•	T	/ o – o
D•	$D_{2,4,5}$	/ {T, D} – {T, D} / S – D
S•	$S_{1,4}$	/ T – T / S – D
D_4	$a(T_3)$	/ o – o
T_3	$a(D_4)$	/ o – {T, D}
H^X•	E	/ o – o
$V_{2,3,4,5,6}$	H^X	/ o – H^X
H^X•	Dg	/ o – o
H^X•	W	/ o – o

$a(T_3)$ bzw. $a(D_4)$ bedeutet, daß es sich hierbei um einen Auftakt handelt, d.h. diese Ersetzung ist nur dann möglich, wenn das Vordersatzanschlußglied mit Auftakt beginnt. Das Zeichen ♩ steht für die Viertelpause.

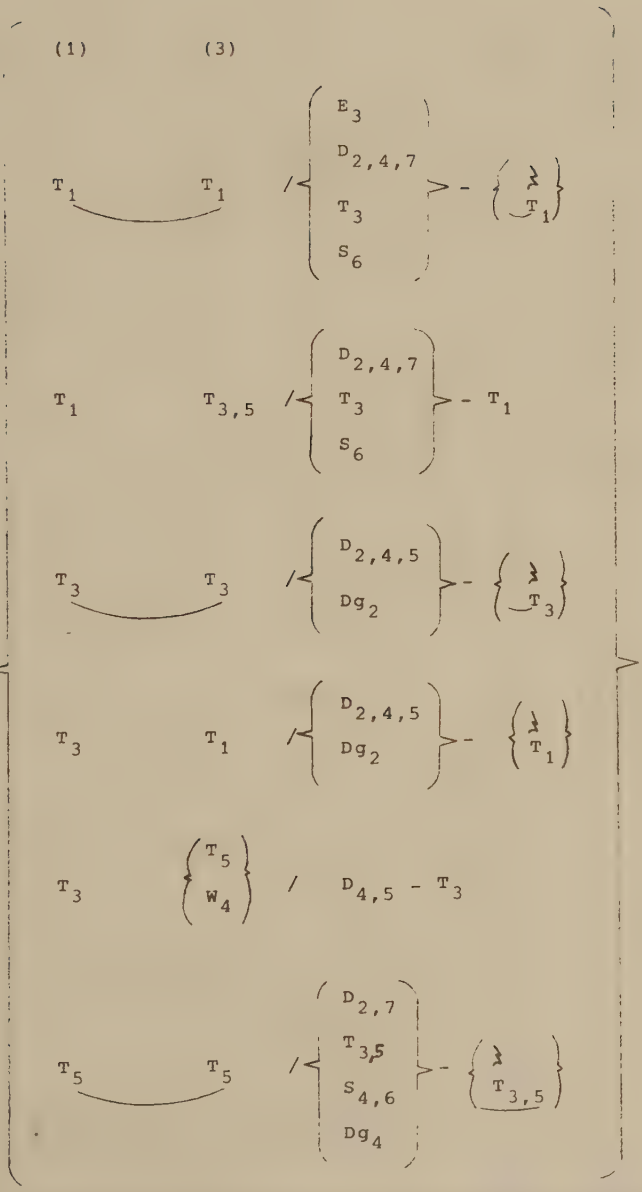
28. $VAT_1F^1_{(),(1),(2)}\rightarrow$

(1)	(3)	
T•	T	/ $\left\{\begin{matrix}T\\D\end{matrix}\right\}$ - o
		/ S - D
D•	D	/ $\left\{\begin{matrix}T\\D\end{matrix}\right\}$ - $\left\{\begin{matrix}T\\D\end{matrix}\right\}$
		/ S - D
S•	S	/ T - o
$V_{2,3,4,5,6}$	H^x	/ o - H^x
H^x•	W	/ o - H^x
H^x•	Dg	/ o - o
T•	D	/ o - $\left\{\begin{matrix}T\\D\end{matrix}\right\}$
D•	T	/ o - o
S•	T	/ T - o

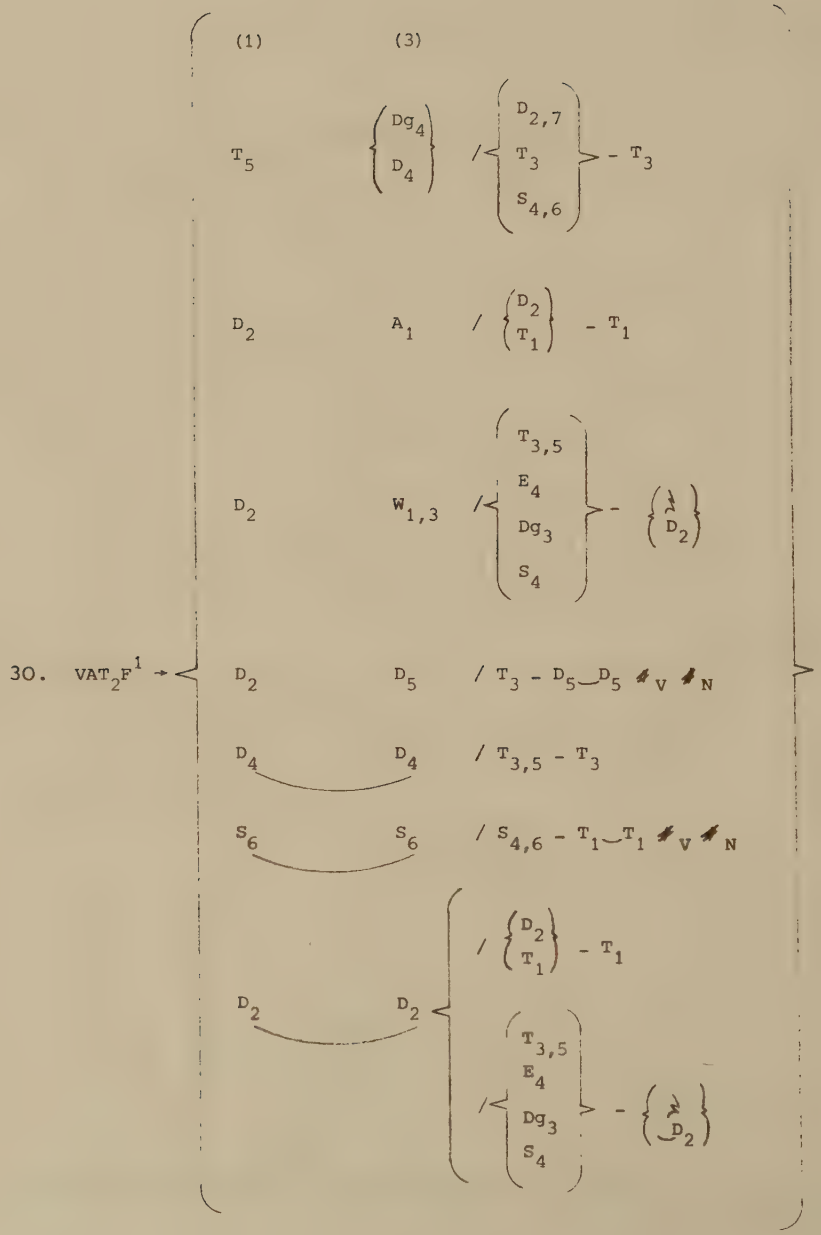
29. $VAT_1F^2_{(),(1),(2)}$

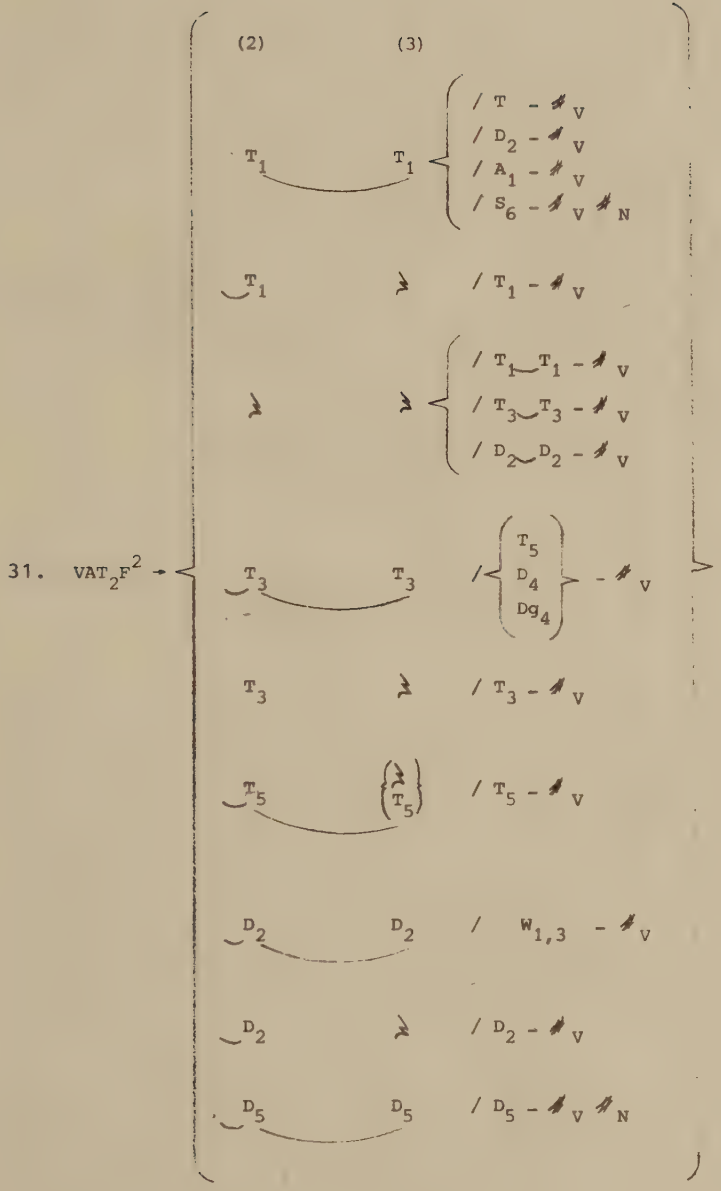
(2)	(3)	
$T\bullet$	T	/ T - o
		/ $\begin{Bmatrix} D \\ S \end{Bmatrix}$ - $\begin{Bmatrix} T \\ D \end{Bmatrix}$
$D\bullet$	D	/ o - $\begin{Bmatrix} T \\ D \end{Bmatrix}$
$S\bullet_{4,6}$	$S_{4,6}$	/ T - o
		/ S - $\begin{Bmatrix} T \\ D \end{Bmatrix}$
$T\bullet$	D	/ o - $\begin{Bmatrix} T \\ D \end{Bmatrix}$
$D\bullet$	T	/ o - o
$S\bullet$	T	/ $\begin{Bmatrix} T \\ S \end{Bmatrix}$ - o
$V_{2,3,4,5,6}$	H^X	/ o - H^X
$H^X\bullet$	Dg	/ o - o
$H^X\bullet$	A_1	/ $\begin{Bmatrix} D \\ S \end{Bmatrix}$ - T_1
$H^X\bullet$	A_3	/ $\begin{Bmatrix} D \\ S \end{Bmatrix}$ - T_3
$H^X\bullet$	E	/ o - o
$H^X\bullet$	W	/ o - o

30. VAT_2F^1 →



Die Fortsetzung der Regel befindet sich auf der nächsten Seite!

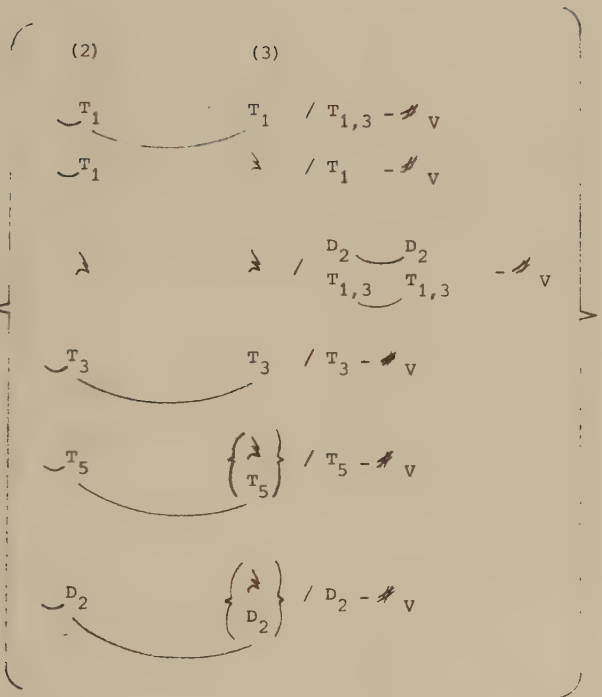

(1)
(3)
T5
Dg4
D4
D2,7
T3
S4,6
- T3
D2
A1
D2
T1
- T1
D2
W1,3
T3,5
E4
Dg3
S4
D2
30. VAT2F1
D2
D5
/ T3 - D5 D5 #V #N
D4
D4
/ T3,5 - T3
S6
S6
/ S4,6 - T1 T1 #V #N
D2
D2
/ D2 T1 - T1
T3,5
E4
Dg3
S4
D2


31. VAT2F2
(2)
(3)
T1
T1
/ T - # V
/ D2 - # V
/ A1 - # V
/ S6 - # V # N
T1
/ T1 - # V
/ T1 T1 - # V
/ T3 T3 - # V
/ D2 D2 - # V
T3
T3
/ T5 D4 Dg4 - # V
T3
/ T3 - # V
T5
T5
/ T5 - # V
D2
D2
/ W1,3 - # V
D2
/ D2 - # V
D5
D5
/ D5 - # V # N

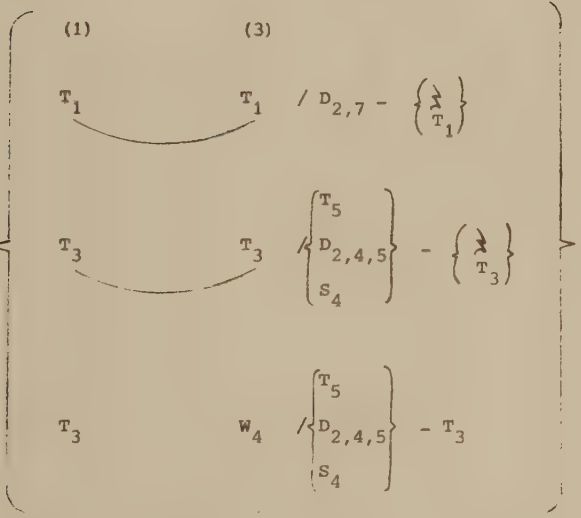
32. $V_1AT_2F^1 \rightarrow$

(1)	(3)
T_1	$T_1 \;/\; D_{2,7} - \left\{\begin{matrix}\lambda \\ T_1\end{matrix}\right\}$
T_1	$T_{3,5} \;/\; D_{2,7} - T_5$
T_3	$T_3 \;/\; \left\{\begin{matrix}T_5 \\ D_{2,4,5} \\ S_4\end{matrix}\right\} - \left\{\begin{matrix}\lambda \\ T_{1,3}\end{matrix}\right\}$
T_5	$\left\{\begin{matrix}D_5 \\ T_5\end{matrix}\right\} \;/\; \left\{\begin{matrix}T_3 \\ Dg_4\end{matrix}\right\} - T_5$
D_2	$\left\{\begin{matrix}D_2 \\ W_{1,3}\end{matrix}\right\} \;/\; \left\{\begin{matrix}S_4 \\ Dg_3\end{matrix}\right\} - D_2$
T_3	$W_4 \;/\; \left\{\begin{matrix}E_6 \\ T_5 \\ D_{2,4,5} \\ S_4\end{matrix}\right\} - T_3$

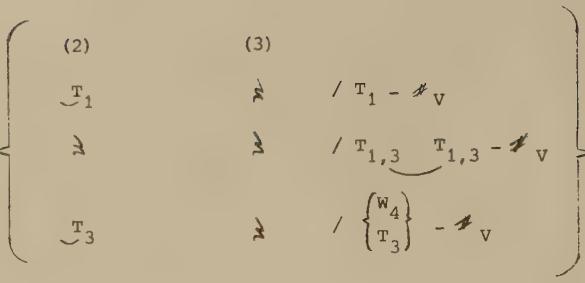
33. $V_1AT_2F^2 \rightarrow$



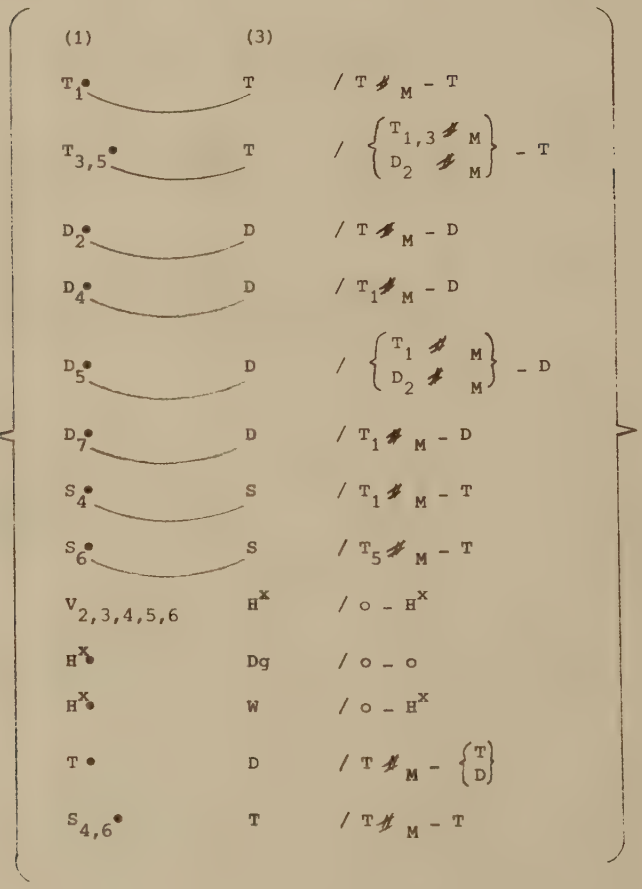
34. $V_2AT_2F^1 \rightarrow$



35. $V_2AT_2F^2 \rightarrow$



36. $MMT_1F^1 \rightarrow$



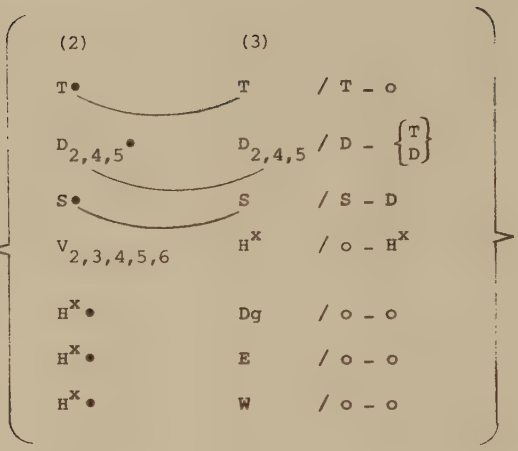
37. $MMT_1F^2 \rightarrow$

(2)	(3)	
T•	T	/ T - o
D•	D	/ D - $\left\{\begin{matrix}T\\D\end{matrix}\right\}$
S•	S	/ $\left\{\begin{matrix}T\\S\end{matrix}\right\}$ - T
S•	T	/ T - T
H^x•	Dg	/ o - o
$V_{2,3,4,5,6}$	H^x	/ o - H^x
H^x•	E	/ o - O
T•	D	/ T - $\left\{\begin{matrix}T\\D\end{matrix}\right\}$
H^x•	W	/ o - o

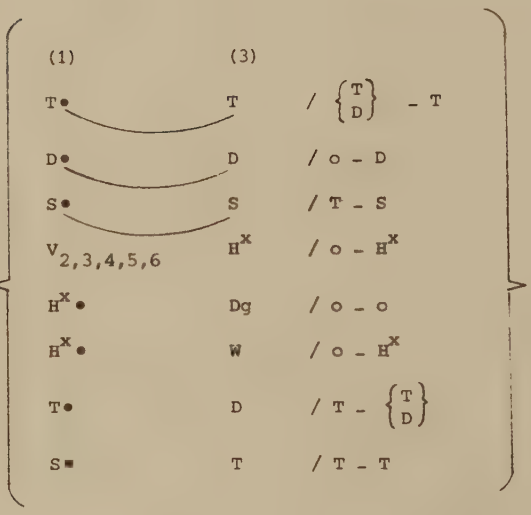
38. $MMT_2F^1 \rightarrow$

(1)	(3)	
T•	T	/ $\left\{\begin{matrix}T\\D\end{matrix}\right\}$ - T
D•	D	/ $\left\{\begin{matrix}T\\D\end{matrix}\right\}$ - D
S•	S	/ T - S
$V_{2,3,4,5,6}$	H^x	/ o - H^x
H^x•	Dg	/ o - o
H^x•	W	/ o - H^x

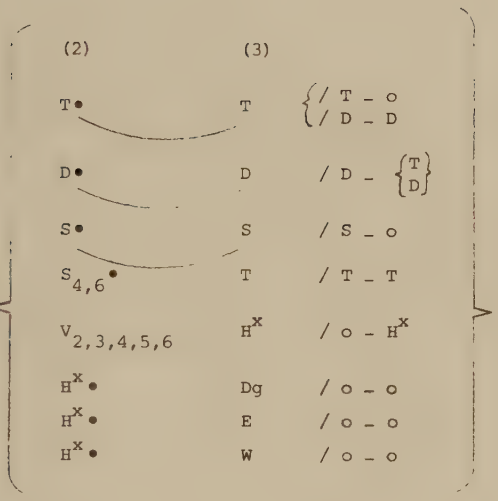
39. $MMT_2F^2 \rightarrow$



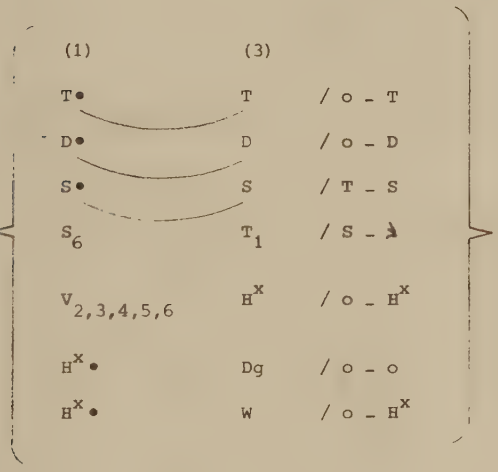
40. $MAT_1F^1 \rightarrow$



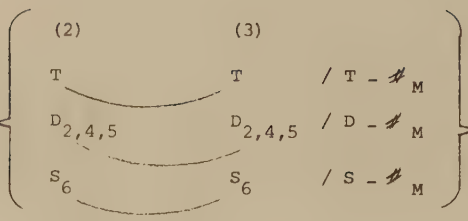
41. MAT_1F^2 →



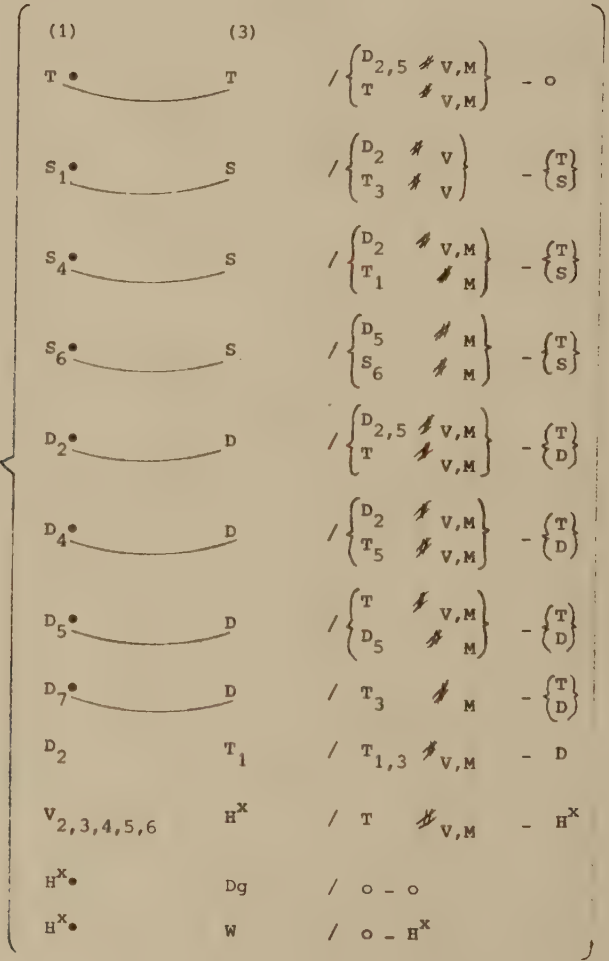
42. MAT_2F^1 →



43. $MAT_2F^2 \rightarrow$

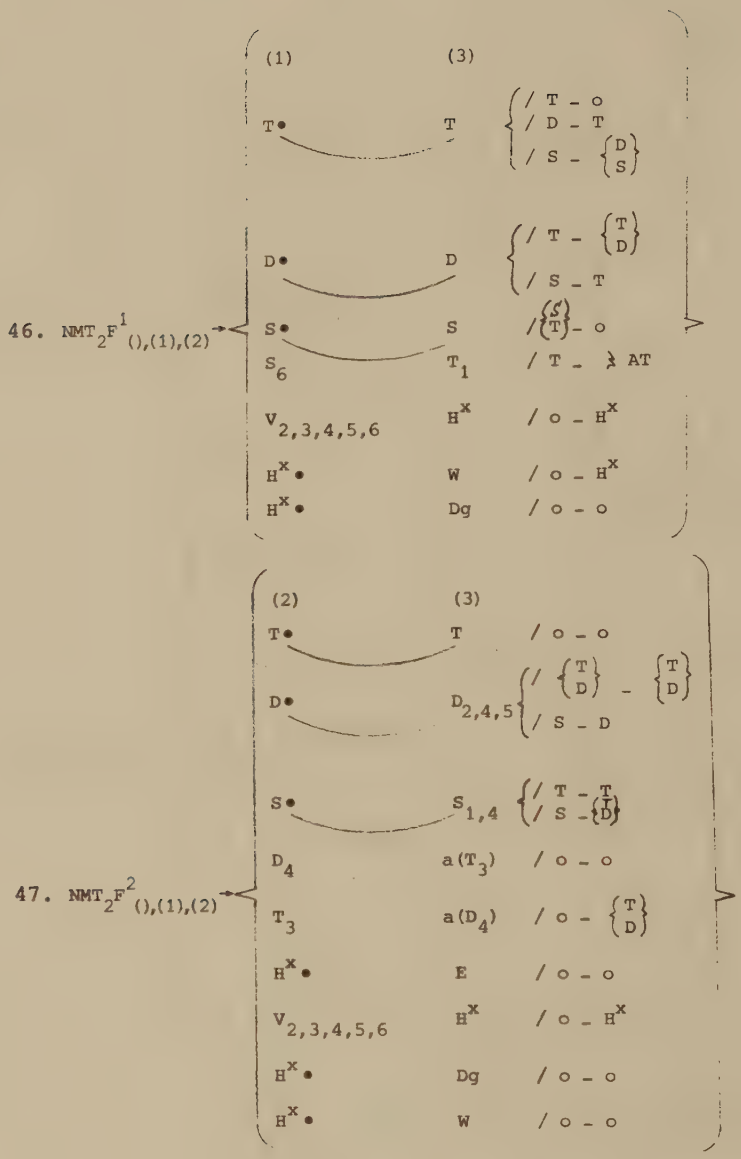


44. NMT_1F^1 (), (1),(2) →

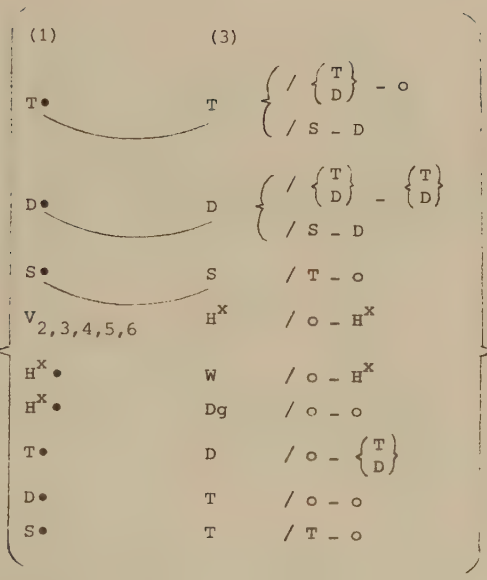


45. $NMT_1F^2_{(),(1),(2)}$ →

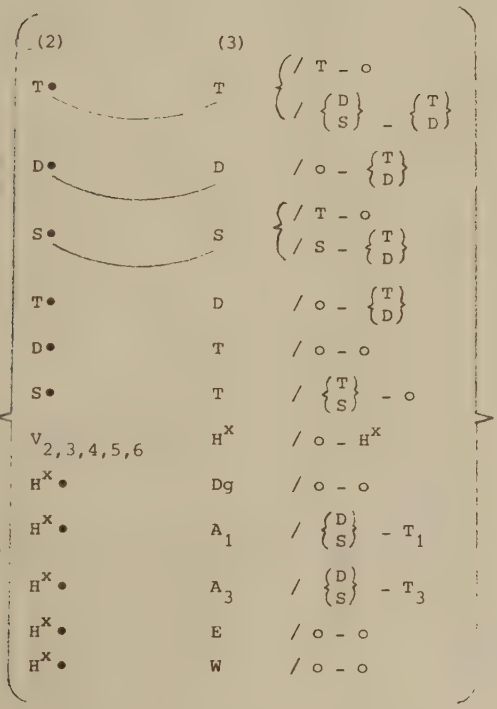
(2)	(3)	
T•	T	/ T - o / S - S
D•	D	/ D - T
S•	S	/ T - {T, D} / S - {S, T}
T•	D	/ T - D
D•	T	/ D - {T, D}
S•	T	/ T - {T, D}
$V_{2,3,4,5,6}$	H^x	/ o - H^x
H^x•	Dg	/ o - o
H^x•	E	/ o - o
H^x•	W	/ o - o


(1)
(3)
T•
T
/ T - o
/ D - T
/ S - {D, S}
D•
D
/ T - {T, D}
/ S - T
46. NMT2F1(),(1),(2)
S•
S
/ {S, T} - o
S6
T1
/ T - ≥ AT
V2,3,4,5,6
H^X
/ o - H^X
H^X•
W
/ o - H^X
H^X•
Dg
/ o - o
(2)
(3)
T•
T
/ o - o
D•
D2,4,5
/ {T, D} - {T, D}
/ S - D
S•
S1,4
/ T - T
/ S - {S, D}
D4
a(T3)
/ o - o
47. NMT2F2(),(1),(2)
T3
a(D4)
/ o - {T, D}
H^X•
E
/ o - o
V2,3,4,5,6
H^X
/ o - H^X
H^X•
Dg
/ o - o
H^X•
W
/ o - o

48. $NAT_1F^1_{(),(1),(2)}$



49. $NAT_1F^2_{(),(1),(2)}$



50. $NAT_2F^1_{(1),(3),(4)}$

(1) (3)

T_1 — $\begin{Bmatrix} \lambda \\ T_{1,3} \\ D_7 \end{Bmatrix}$ $/ \begin{Bmatrix} T_5 \\ D_{2,5,7} \\ Dg_2 \\ E_3 \end{Bmatrix}$ $- \begin{Bmatrix} T_1 \\ \lambda \end{Bmatrix}$

T_3 — T_3 $\begin{cases} / D_{4,5} - \begin{Bmatrix} \lambda \\ T_3 \end{Bmatrix} \\ / D_2 - T_5 \end{cases}$

$\begin{Bmatrix} T_3 \\ D_{2,4} \\ D_{2,5,7} \end{Bmatrix}$ — $\begin{matrix} D_{2,7} \\ D_7 \end{matrix}$ $/ \begin{Bmatrix} T_{1,3} \\ D_{2,7} \\ S_4 \\ Dg_4 \end{Bmatrix} - T_1$

T_5 — T_5 $/ \begin{Bmatrix} S_{4,6} \\ Dg_4 \\ E_6 \end{Bmatrix} - T_{3,5}$

T_5 W_6 $/ \begin{Bmatrix} S_{4,6} \\ Dg_4 \\ E_6 \end{Bmatrix} - T_{5,3}$

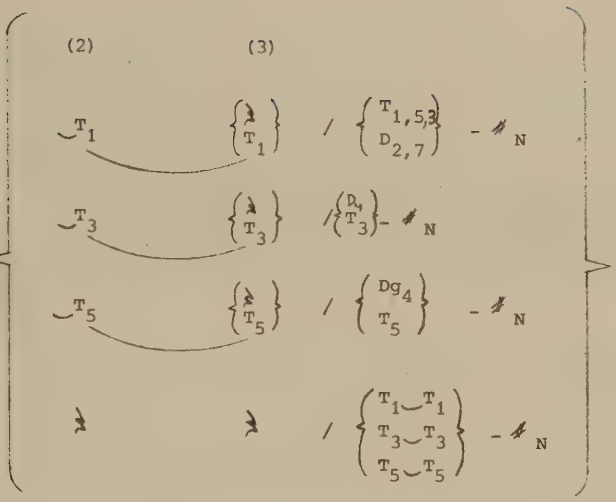
T_5 $\begin{Bmatrix} D_4 \\ Dg_4 \end{Bmatrix}$ $/ \begin{Bmatrix} S_{4,6} \\ Dg_6 \end{Bmatrix} - T_3$

T_3 $\begin{Bmatrix} Dg_4 \\ D_4 \end{Bmatrix}$ $/ D_{2,4,5} - T_5$

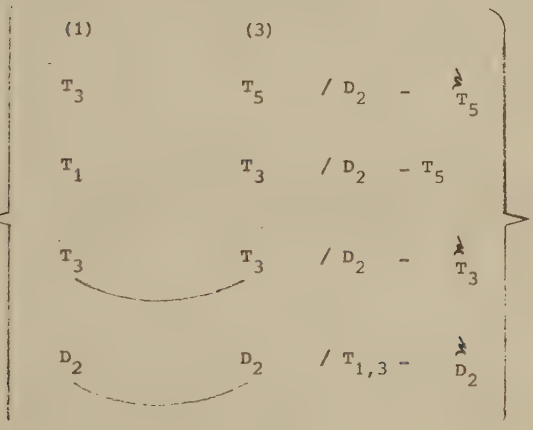
S_6 D_7 $/ T_{1,5} - T_1$

D_7 T_3 $/ D_{2,4,7} - T_1$

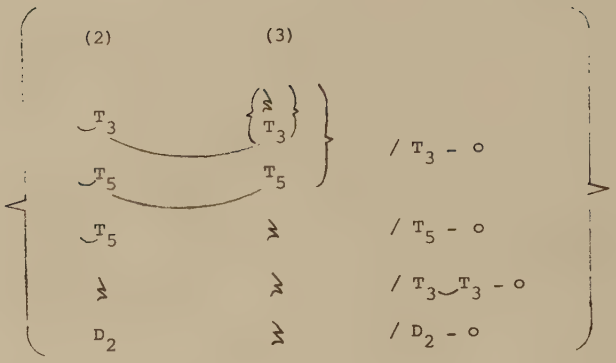
51. $NAT_2F^2 \rightarrow$



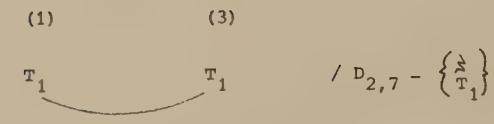
52. $N_1AT_2F^1 \rightarrow$



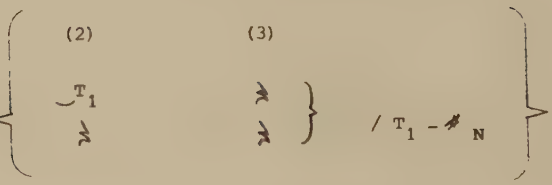
53. $N_1AT_2F^2 \rightarrow$



54. $N_2AT_2F^1 \rightarrow$



55. $N_2AT_2F^2 \rightarrow$



1.2.4 T - Regeln der Syntax

1. Auftakttransformationen:

Bei der Generierung eines Auftaktes zum Vordersatz handelt es sich um eine Tilgung der letzten Zählzeit aus dem Satz $N(NAT_2F^2Z^3)$ und um eine Adjunktion des Auftaktgliedes AT vor die erste Zählzeit des Vordersatzes $(VMT_1F^1Z^1)$:

$$a(V_{(),(1),(2)}): SB_a = \left[\#_V VMT_1F^1Z^1,\ldots,VAT_2F^2Z^3\#_V, \#_N NMT_1F^1Z^1,\ldots, NAT_2F^2Z^3\#_N\right]$$

$$\Longrightarrow SV_a = \left[\#_V AT,VMT_1F^1Z^1,\ldots,VAT_2F^2Z^3\#_V, \#_N NMT_1F^1Z^1,\ldots,NAT_2F^2Z^2\#_N\right]$$

Bei allen weiteren Auftakttransformationen handelt es sich um eine Substitution der Zählzeit mit Betonungsrang (3) im zweiten Fuß des Taktes, der demjenigen Takt vorausgeht, zu dem ein Auftakt gebildet werden soll:

$$a(Y): \quad SB_a = \left[_\ XF^2Z^3, _\ YF^1Z^1\right] \Longrightarrow SV_a = \left[\ AT, _\ YF^1Z^1\ \right]$$

X bezeichnet den Namen der Konstituentenklasse, die dem Konstituenten vorausgeht, zu dem der Auftakt gebildet werden soll (Konstituent Y).

2. Pausentransformationsregeln:

Obligatorische Pausen wurden bereits durch die Ersetzungsregeln eingeführt, optionale Pausen werden nun durch Substitutionstransformationen eingeführt (prinzipiell wäre dies auch durch Ersetzungsregeln möglich). Die Indizes an den eckigen Klammern geben an, für welche Konstituentenklasse die Regel jeweils gilt.

a) $SB_{\nearrow} = \begin{bmatrix} 1.: H, H, H, H \\ 2.: H, H, H, AT \end{bmatrix}_{VMT_{1(),(1),(2)}} \Longrightarrow$

$SV_{\nearrow} = \begin{bmatrix} 1.: H \nearrow H \nearrow \\ 2.: H, H \nearrow AT \end{bmatrix}_{VMT_{1(),(1),(2)}}$

b) $SB_{\nearrow} = \begin{bmatrix} 1.: H, H, H, H \\ 2.: H, H, H, AT \end{bmatrix} \begin{pmatrix} VMT_{2(),(1),(2)} \\ MMT_{2} \\ NMT_{2(),(1),(2)} \end{pmatrix} \Longrightarrow$

$SV_{\nearrow} = \begin{bmatrix} 1.: H \nearrow \nearrow \nearrow \\ 1.: H, H \nearrow \nearrow \\ 2.: H, H \nearrow AT \end{bmatrix} \begin{pmatrix} VMT_{2(),(1),(2)} \\ MMT_{2} \\ NMT_{2} \end{pmatrix}$

$SV_{\nearrow} = \begin{bmatrix} 1.: H, H \nearrow \nearrow \\ 2.: H, H \nearrow AT \end{bmatrix}_{MMT_{2}}$

3. $w\approx$ Transformation:

Die $w\approx$ Transformation stellt eine Wiederholung eines Satzes her, in dessen letztem Takt Änderungen vorgenommen werden, damit dieser der Harmoniedistribution des letzten Taktes des Satzes entspricht, für den der wiederholte Satz stehen soll. Wird z.B. der Vordersatz als Nachsatz wiederholt, dann muß die Harmoniefolge des letzten Taktes des wiederholten Vordersatzes den Möglichkeiten der Harmoniefolge im Nachsatz angepaßt werden, was für die vorausgehenden Takte nicht notwendig ist, da dort die Distribution der Harmoniefolgen für Vordersatz und Nachsatz gleich ist.

$$w\approx (V)_N : \; SB_w = \left[VMT_1, \; VMT_2, \; VAT_1, \; VAT_2 \right]_N \Longrightarrow$$

$$SV_w = \left[VMT_1, \; VMT_2, \; VAT_1, \; NAT_2 \right]_N$$

Analog wird die Substitution des letzten Taktes bei den Transformationen $w\approx (V_1)_{V_2}$, $w\approx (V_2)_N$ und $w\approx (N_1)_{N_2}$ vorgenommen.

4. $w=$ Transformation:

Die $w=$ Transformation ist eine identische Wiederholung einer Konstituentenklasse X als Vertreter für eine andere Konstituentenklasse Y :

$$w= (X)_Y : \; SB_{w=} \left[X \right]_X \Longrightarrow SV_{w=} = \left[X \right]_Y$$

5. s - Transformation:

Die s - Transformation besteht in einer ordinalen Verschiebung einer Tonfolge im Bereich von einer Sekunde bis zu einer Quinte aufwärts oder abwärts:

$$s(X):\ SB_s = \left[z_i^1 , z_j^2 , z_k^3 , z_l^4 \right]_{4/4} \Longrightarrow SV_s = \left[z_{i\pm s}^1 , z_{j\pm s}^2 , z_{k\pm s}^3 , z_{l\pm s}^4 \right]_{4/4}$$

wobei z für eine Zählzeit steht, die Indizes i, j, k und l die Tonstufe angeben und s das Verschiebungsintervall darstellt. Dies ist hier für den 4/4 Takt dargestellt, für die anderen Metren ist die formale Darstellung analog zu schreiben.

6. u - Transformation:

$$u(X):\ SB_u = \left[z_i^1 , z_j^3 , z_k^2 , z_l^3 \right]_{4/4} \Longrightarrow SV_u = \left[z_l^1 , z_k^3 , z_j^2 , z_i^3 \right]_{4/4}$$

Wir haben hier das Beispiel der Umkehrung eines ganzen Taktes im 4/4 - Metrum gewählt, für andere Konstituentenklassen ist die Transformation formal analog zu schreiben. s - und u - Transformationen sind auf demselben hierarchischen Niveau gemeinsam anwendbar.

7. Punktierungstransformation:

$$SB_\bullet = \left[1\, z^1,\ 1\, z^3 \right] \Longrightarrow SV_\bullet = \left[1\, \tfrac{1}{2}\, z^1,\ \tfrac{1}{2}\, z^3 \right]$$

Die Zahl vor den Zählzeitsymbolen z^1 und z^3 stellt die relative Zeitdauer der Notenwerte auf diesen Zählzeiten dar.

8. Bindungstransformation:

$$SB_{\smile} = \left[z^1, z^3 \right] \Longrightarrow SV_{\smile} = \left[z^1 + z^3 \right]$$

Das Pluszeichen zwischen den beiden Zählzeitsymbolen soll anzeigen, daß es sich nicht um getrennte Zählzeiten, sondern um miteinander verbundene Zählzeiten handelt.

1.2.5 Offene Probleme

Entgegen unserer ursprünglichen Intention haben wir einige Pausentransformationen eingeführt, obwohl diese auch durch Ersetzungsregeln darstellbar sind (vgl. z.B. Regel 35.). Der Grund hierfür besteht darin, daß die Anzahl der Lieder, die innerhalb der ersten drei Takte einer Phrase (V, M oder N) Pausen enthalten, in unserem Corpus sehr klein war, so daß eine Verallgemeinerung der Pausenregeln noch auf sehr schwachen Füßen steht. Da wir ohnehin nur Melodien für musikpsychologische Experimente zu generieren beabsichtigen, in denen (außer am Ende des Nachsatzes) keine Pausen vorkommen sollen, um Segmentierungsprozesse unabhängig von derartig markanten Indikatoren für Segmentgrenzen, wie sie Pausen darstellen, untersuchen zu können, hat dieser Mangel keine Konsequenzen.

Einigermaßen stiefmütterlich wurden in unseren Distributionsanalysen rhythmische Abhängigkeiten auf Niveaus oberhalb eines Taktes behandelt, weil wir primär erst einmal an der Wahrnehmung melodisch-harmonischer Strukturen interessiert sind und deshalb die rhythmische Struktur bei den in Experimenten zu verwendenden generierten Melodien konstant halten werden. Dieser Mangel ist aber mit prinzipiell denselben Methoden der Distributionsanalyse zu beheben.

Bei der harmonischen Analyse wurden Subdominantparallelen, Dominantparallelen und Dominanten der Dominante wie einfache Subdominanten bzw. Dominanten behandelt. Bei einer differenzierteren harmonischen Analyse, die diese Besonderheiten mit berücksichtigt, wäre die Syntax allgemeiner geworden, was aber wegen der angestrebten Einfachheit der mit Hilfe dieser Syntax zu generierenden Lieder für Experimente nicht unbedingt einen gravierenden Mangel im Rahmen dieser Verwendung darzustellen braucht.

Für den Musikwissenschaftler mag die vorliegende Syntax noch unbefriedigend sein, für den Musikpsychologen stellt sie aber eine verwendbare Syntax dar, die eine echte Teilmenge aller als syntaktisch richtig zu betrachtenden Volks- und Kinderliedmelodien generieren kann. Hier eröffnet sich für den Musikwissenschaftler noch ein breites Arbeitsfeld.

2. Beispiel für die Generierung einer Melodie

Als Beispiel für die generative Anwendung der Syntax wollen wir eine Melodie ableiten, die aus drei Phrasen bestehen soll, in der Tonart C-Dur geschrieben sein und einen 4/4 - Takt besitzen soll, in dem keine Auftakte vorkommen. Außerdem wollen wir festlegen, daß nur Viertelnoten und höchstens am Ende einer Phrase halbe Noten vorkommen sollen.

Die Ableitung beginnt mit Regel 1, wo wir uns für die erste Zeile entscheiden:

1. $P \rightarrow \not\!\!\!

Das Anschlußglied wird durch Regel 16 ersetzt, aus der wir die 1. Zeile gewählt haben:

16. $VA \rightarrow VAT_1 \quad VAT_2$

Nun folgt das Mittelsatzmotiv, das durch Regel 18 ersetzt wird; auch hier wählten wir die 1. Zeile:

18. $MM \rightarrow MMT_1 \quad MMT_2$

Da das Mittelsatzanschlußglied nach Regel 14 durch eine Wiederholung des Motivs generiert wird, können wir jetzt schon zur Ersetzung des Nachsatzmotivs übergehen. Hierfür benötigen wir die Regel 20, aus der wir die 2. Zeile wählen:

20. $NM \rightarrow NMT_1 \quad s(NMT_1)NMT_2$

Das Anschlußglied bilden wir durch die 1. Zeile aus Regel 22:

22. $NA \rightarrow NAT_1 \quad NAT_2$

Damit sind alle Motive und Anschlußglieder durch ihre Takte ersetzt. Der nächste Ableitungsschritt besteht in der Anwendung der Regel 23, durch die alle Takte durch Füße mit den beiden Hauptbetonungen ersetzt werden:

23. $_ T \rightarrow _ F^1 \quad _ F^2$

Nun beginnt die Ersetzung der Füße durch die terminalen Elemente (Tonstufen mit harmonischen Funktionen). Bei diesen Ersetzungen müssen immer die Metaregeln für Kontexte zur Einschränkung der Tonkombinationen berücksichtigt werden.

Der erste Fuß des Vordersatzmotivs im 1. Takt wird nach Regel 24 ersetzt. Hier wählen wir die 1. Zeile mit den Tonstufen 3 und 1:

		(1)	(3)
24.	VMT_1F^1	→ T_3	T_1

Da die Betonungsränge, die durch die Ersetzungsregeln generiert werden, in allen Takten gleich sind, werden wir sie nur noch in der nächsten Regel angeben.

Der 2. Fuß des 1. Taktes des Vordersatzmotivs wird durch Regel 25 ersetzt, wobei wir die 1. Zeile verwenden und die 1. und 5. Tonstufe auswählen:

		(2)	(3)
25.	VMT_1F^2	→ T_1	T_5

Der 2. Takt wird durch die Umkehrungstransformation aus dem 1. Takt nach Regel 15 durch die u - Transformationsregel abgeleitet:

$$u\,(VMT_1F^1,\ VMT_1F^2) \rightarrow (T_5,\ T_1,\ T_1,\ T_3)$$

Die Füße des Vordersatzanschlußgliedes ersetzen wir wie folgt:

28. $VAT_1F^1 \rightarrow D_4\ D_5$ (2. Zeile der Regel)

29. $VAT_1F^2 \rightarrow T_3\ T_1$ (1. Zeile der Regel)

30. $VAT_2F^1 \rightarrow D_2\ A_1$ (8. Zeile der Regel)

31. $VAT_2F^2 \rightarrow T_1\ T_1$ (1. Zeile der Regel)

Die Füße des Mittelsatzmotivs ersetzen wir wie folgt:

36. $MMT_1F^1 \rightarrow D_2 \quad D_2$ (2. Zeile der Regel)

37. $MMT_1F^2 \rightarrow D_4 \quad D_2$ (2. Zeile der Regel)

38. $MMT_1F^1 \rightarrow D_5 \quad D_4$ (2. Zeile der Regel)

39. $MMT_2F^2 \rightarrow T_3 \quad T_3$ (2. Zeile der Regel)

Da das Mittelsatzanschlußglied eine Wiederholung des Mittelsatzmotivs ist (vgl. Regel 14), ist die Ersetzung der Füße des Anschlußgliedes gleich der des Motivs, so daß wir gleich zum Nachsatz übergehen können.

Die Füße des Nachsatzmotivs ersetzen wir wie folgt:

44. $NMT_1F^1 \rightarrow T_3 \quad Dg_4$ (1. und 6. Zeile der Regel)

45. $NMT_1F^2 \rightarrow T_5 \quad T_5$ (1. Zeile der Regel)

Da der Takt 2 des Nachsatzmotivs eine Sequenz des 1. Taktes ist, wenden wir die s - Transformation an und wählen als Intervall eine Sekunde abwärts:

$s_{-1} (NMT_1F^1, NMT_1F^2) \rightarrow (D_2, Dg_3, D_4, D_4)$

Die Füße des Nachsatzanschlußglieds ersetzen wir wie folgt:

48. $NAT_1F^1 \rightarrow T_1 \; Dg_2$ (1. und 6. Zeile der Regel)

49. $NAT_1F^2 \rightarrow T_3 \; W_4$ (1. und letzte Zeile der Regel)

50. $NAT_2F^1 \rightarrow T_3 \; D_2$ (3. Zeile der Regel)

51. $NAT_2F^2 \rightarrow T_1 \frown T_1$ (1. Zeile der Regel)

Damit sind alle nonterminalen Symbole durch terminale Elemente ersetzt und die Generierung der Melodie abgeschlossen. Der Notentext dieser Melodie mit ihrem Strukturbaum ist in Abb. 2 dargestellt.

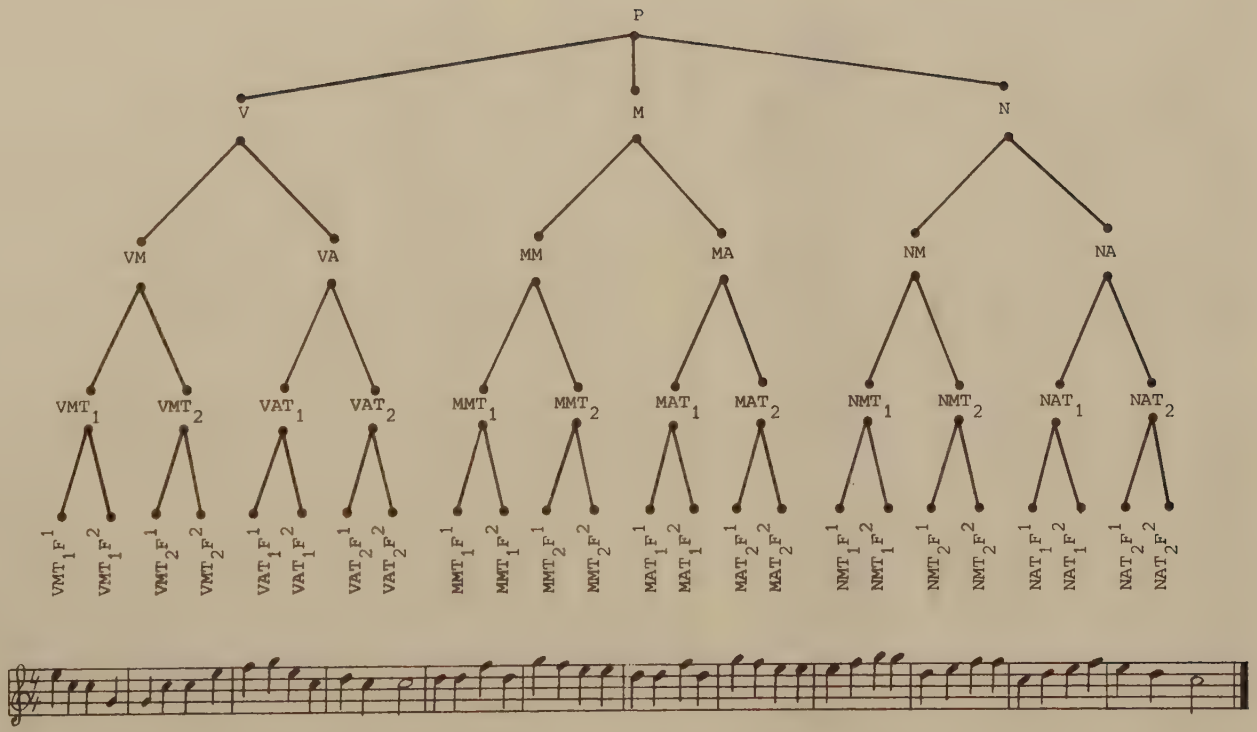


Abb. 2.: Konstituentenstruktur und Notentext der generierten Melodie

Anhang Lieder des Corpus
(hinter dem Titel steht in Klammern die Quelle)

1. Wildgänse rauschen durch die Nacht (Klusen)
2. Es geht eine helle Flöte (Klusen)
3. Wenn wir erklimmen (Klusen)
4. Winde wehn, Schiffe gehn (Schüngeler)
5. Bald gras ich am Neckar (Lutz)
6. Und in dem Schneegebirge (Klusen)
7. Die güldene Sonne (Schüngeler)
8. Wenn ich ein Vöglein wär (Lutz)
9. Schneeflöckchen, Weißröckchen (Lutz)
10. Kommt ein Vogel geflogen (Lutz)
11. Ein Vogel wollte Hochzeit machen (Klusen)
12. Es war einmal ein kleiner Mann (Lutz)
13. Alle meine Entchen (Lutz)
14. Taler, Taler (Lutz)
15. Mariechen saß auf einem Stein (Lutz)
16. Abend wird es wieder (Schüngeler)
17. Der Mond ist aufgegangen (Schüngeler)
18. Dort auf jenem Berge (Lutz)
19. Heute ist im Schloß ein großer Ball (Lutz)
20. Wer hat die schönsten Schäfchen (Ziegenrücker)
21. Auf dem grünen Rasen (Lutz)
22. Ade zur guten Nacht (Schüngeler)
23. Laterne, Laterne (Lutz)
24. Die Blümlein, sie schlafen (Schüngeler)
25. Sah ein Knab ein Röslein stehn (Schüngeler)
26. Stille Nacht, heilige Nacht (Lutz)
27. Vöglein im hohen Baum (Schüngeler)
28. Winter ade (Lutz)
29. Kuckuck, Kuckuck ruft's aus dem Wald (Lutz)
30. Das Wandern ist des Müllers Lust (Klusen)
31. Gestern abend ging ich aus (Lutz)
32. Trairiro, der Sommer, der ist do! (Lutz)
33. A, B, C, die Katze lief in'n Schnee (Lutz)

34. Häschen in der Grube (Lutz)
35. Das Lieben bringt groß Freud (Schüngeler)
36. Alle Vögel sind schon da (Lutz)
37. Hopp, hopp, hopp (Lutz)
38. Summ, summ, summ (Lutz)
39. Hänschen klein (Lutz)
40. Suse, liebe Suse (Lutz)
41. Tanz rüber, tanz nüber (Lutz)
42. Heut soll das große Flachsernten sein (Schüngeler)
43. Ich weiß nicht, was soll das bedeuten (Schüngeler)
44. Jetzt kommen die lustigen Tage (Schüngeler)
45. Im schönsten Wiesengrunde (Schüngeler)
46. Hänsel und Gretel (Lutz)
47. Weißt du, wieviel Sternlein stehen? (Lutz)
48. Guter Mond (Lutz)
49. Ihr Kinderlein kommet (Lutz)
50. Ein Männlein steht im Walde (Lutz)
51. Im Märzen der Bauer (Lutz)
52. Es tanzt ein Bi-Ba-Butzemann (Lutz)
53. Es klappert die Mühle am rauschenden Bach (Lutz)
54. Wenn alle Brünnlein fließen (Schüngeler)
55. Es wollt ein Schneider wandern (Schüngeler)
56. Drunten im Unterland (Schüngeler)
57. Ich wollt ein Bäumlein steigen (Schüngeler)
58. Der Winter ist vergangen (Schüngeler)
59. Muß i denn zum Städtele naus (Klusen)
60. Heißa, Kathreinerle (Schüngeler)
61. Nun will der Lenz uns grüßen (Schüngeler)
62. Im Märzen der Bauer (Schüngeler)
63. Die Gedanken sind frei (Klusen)
64. Sabinchen war ein Frauenzimmer (Klusen)
65. Nun ade, du mein lieb Heimatland (Klusen)
66. Dat du min Leevsten büst (Ziegenrücker)
67. Fuchs, du hast die Gans gestohlen (Lutz)
68. In einem kühlen Grunde (Schüngeler)
69. Am Brunnen vor dem Tore (Klusen)
70. Der Kuckuck und der Esel (Lutz)
71. Jetzt gang i ans Brünnele (Schüngeler)

A N H A N G 2

S y n t a x t r a i n i n g

Trainingstext

Einführung

Ihr nehmt in den nächsten Wochen an Untersuchungen teil, in denen der Einfluß des beim Hörer vorhandenen Wissens über die in Musik vorfindbaren Regelmäßigkeiten des Aufbaus auf die Wahrnehmungsvorgänge beim Hören von Musik untersucht werden soll. Da man von Euch nicht voraussetzen kann, daß Ihr bereits über diese Kenntnisse im Detail verfügt, führt der Versuchsleiter zuerst ein Training mit Euch durch, in dessen Verlauf Ihr Euch einige Kenntnisse über den in Volksliedern vorfindbaren Strukturaufbau aneignen könnt.

Das Gelingen des Trainings hängt auch von Eurer Bereitwilligkeit zur Mitarbeit ab. Deshalb möchten wir Euch bitten, daß Ihr Euch aktiv am Training beteiligt, d.h. wir möchten Euch ermuntern, immer dann Fragen zu stellen, wenn Ihr glaubt, etwas nicht verstanden zu haben, oder wenn im Text Unklarheiten enthalten sein sollten. Es kommt uns nämlich darauf an, daß alle am Ende des Trainings den gleichen Wissensstand haben. Um dieses Ziel erreichen zu können, ist es unbedingt notwendig, daß Ihr zu allen Trainingssitzungen kommt.

Wir werden am Ende einer Sitzung versuchen, uns auf einen möglichst für Euch alle günstigen Termin für die nächste Sitzung zu einigen. Sollte es einem von Euch einmal nicht möglich sein, zu einer Trainingssitzung zu kommen, dann ruft bitte vorher den Projektleiter an (Stoffer, 700-2664) oder schreibt rechtzeitig eine Karte (Th. Stoffer, Psychologisches Institut, Ruhr-Universität Bochum, GAFO 02/382, 4630 Bochum). In diesem Fall achtet bitte auch darauf, daß sofort ein Ersatztermin mit Euch vereinbart wird. Der Versuchsleiter wird so weit wie möglich versuchen, Euren Terminwünschen entgegenzukommen. Wenn Ihr es aber immer einrichten könntet, zu allen Trainingssitzungen zu kommen, würdet Ihr uns eine Menge Zeit und Arbeit ersparen.

Wir bitten darum, daß Ihr während der Trainingssitzungen Euch Notizen macht. Am Ende einer Trainingssitzung schaut bitte Eure Notizen noch einmal durch und überlegt, ob noch irgendwelche Unklarheiten bestehen. Wenn dies der Fall ist, habt Ihr dann noch Gelegenheit, Fragen zu stellen. Am Tag vor der jeweils nächsten Trainingssitzung schaut bitte für ca. eine viertel Stunde in Eure Notizen und wiederholt den Inhalt der letzten Trainingssitzung.

Vielen Dank für Eure Bereitschaft mitzumachen!

Erste Trainingssitzung

Einführung

Ihr alle besitzt, da Ihr ein Instrument spielen gelernt habt, in mehr oder weniger großem Umfang musikalisches Wissen, das sich nicht nur auf die technische Beherrschung des Instruments bezieht, sondern auch auf die Analyse musikalischer Formen und Strukturen. Dieses Wissen dient Euch zur adäquaten Interpretation eines Musikstücks. So könnt Ihr z.B. einzelne Motive voneinander unterscheiden oder Beziehungen zwischen ihnen herstellen, Ihr könnt z.T. Tonartenwechsel erkennen, rhythmische Varianten voneinander unterscheiden, melodische Kontraste bemerken, motivische Verarbeitungen erkennen und dgl. mehr. All dieses mehr oder weniger bewußt verfügbare Wissen dient dazu, ein Musikstück so zu spielen, daß der Hörer die von Euch herausgearbeiteten Teile als zusammengehörig oder voneinander abgehoben erkennen kann, sofern er selbst über entsprechende Kenntnisse und Erfahrungen verfügt.

Dieses Nachvollziehen der musikalischen Struktur durch den Hörer ist Gegenstand eines Teils der Versuche, an denen Ihr im Anschluß an das Training teilnehmen werdet. Ihr werdet Euch nach dem Training in der Situation eines Hörers befinden, der über relativ differenzierte Kenntnisse über musikalische Strukturen verfügt, allerdings nur über die Strukturen, die innerhalb einer ganz begrenzten musikalischen Form anzutreffen sind, nämlich in einer Unterform der deutschen Volks- und Kinderlieder. Im Verlauf des Trainings werden wir Euch mit einigen Details dieser Musik bekanntmachen, und Ihr werdet Gelegenheit haben, Eure Fähigkeit zu üben, ganz bestimmte formale und strukturelle Details zu erkennen.

Das was Ihr nun lernen werdet, ist vergleichbar mit dem Erlernen der Grammatik einer Fremdsprache. Ohne die Beherrschung der Grammatik kann man eine Sprache nicht verstehen, selbst dann nur äußerst unzureichend, wenn man die in der Sprache vorkommenden Wörter und ihre Bedeutung bereits kennt. In ziemlich genau dieser Situation befinden die meisten von Euch sich jetzt: Ihr kennt natürlich die musikalischen Tonleitern, die verschiedenen Metren und Rhythmen, auch einige Harmonien, die in Musik vorkommen können, d.h. Ihr kennt die Elemente, aus denen Musik besteht. Was Ihr aber über die musikalische Form des Volks- und Kinderliedes noch nicht oder nur unvollständig wißt, ist die Art und Weise, wie diese musikalischen Elemente miteinander kombiniert auftreten. Im Verlauf

des Trainings werdet Ihr also einen großen Teil der Grammatik der Volks- und Kinderlieder lernen. Dies wird Euch dazu befähigen, Zusammenhänge und Kontraste zu identifizieren und spezifische Erwartungen über den Fortgang einer gerade zu hörenden Melodie zu bilden.

Bei diesem Training und den anschließenden Versuchen kommt es uns ausschließlich auf das Hören und Identifizieren struktureller Merkmale von Melodien an. Uns ist bewußt, daß dies nicht gleichzusetzen ist mit Musikhören in all seinen vielfältigen Formen. Z.B. bleibt der in Musik manchmal zum Ausdruck gebrachte Stimmungsgehalt dabei völlig unberücksichtigt. Wir möchten Euch deshalb bitten, Euch beim Anhören der im Verlauf des Trainings vorgespielten Beispiele ganz auf die Struktur der Melodie zu konzentrieren und von einem Erfassen des Stimmungsgehalts abzusehen. Dies wird ohnehin dadurch erschwert, daß die meisten zu hörenden Melodien rhythmisch relativ eintönig sind.

Soviel an Vorbemerkungen. Auf der nächsten Seite geht es dann richtig los!

Phrasenstrukturen I: Phrasen und Subphrasen

Ihr werdet im Training nun die Grammatik der sogenannten "zweiteiligen" Volks- und Kinderlieder kennenlernen. Wir wollen zunächst damit beginnen, uns einige Lieder dieser Kategorie anzuhören. Eure Aufgabe ist es dabei, jedesmal die Grenze zwischen den beiden Teilen herauszufinden und durch einen Strich in den Notenlinien zu markieren. Die Beispiele findet Ihr auf S. 1 des Arbeitshefts. Hören wir uns nun das erste Beispiel an.

Wo genau befindet sich die Grenze zwischen den beiden Teilen?

Diese beiden Teile werden Vordersatz und Nachsatz genannt. Bitte sucht nun auch im zweiten Beispiel nach der Grenze zwischen dem Vordersatz und dem Nachsatz.

Wo befindet sich in diesem Beispiel die Grenze zwischen Vorder- und Nachsatz?

Beide Lieder bestehen also aus acht Takten, und in beiden befindet sich die Grenze zwischen Vorder- und Nachsatz genau in der Mitte des Liedes. Ist dies eigentlich immer so? Hören wir uns noch ein weiteres Beispiel an und markiert bitte wieder die Grenze zwischen Vorder- und Nachsatz.

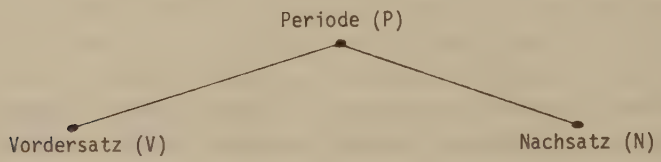
Wo befindet sich in diesem Lied die Grenze zwischen Vorder- und Nachsatz?

Wir werden gleich ein wenig systematisch untersuchen, warum in diesem Fall auch noch der 5. und 6. Takt zum Vordersatz zu rechnen ist und nicht zum Nachsatz, doch zuvor wollen wir uns noch das Beispiel 4 anhören und die Grenze zwischen Vordersatz und Nachsatz markieren.

Wo liegt in diesem Beispiel die Grenze zwischen Vorder- und Nachsatz?

Nun wollen wir beginnen, die Struktur solcher zweiteiligen Volks- und Kinderlieder ein wenig systematisch zu behandeln.

Eine ganze Liedstrophe bezeichnen wir als musikalische Periode. Die Periode wird bei den zweiteiligen Liedern, wie es der Name schon sagt, in zwei Teile untergliedert, nämlich in den am Anfang stehenden Vordersatz und den am Ende stehenden Nachsatz. Einteilige Lieder bestehen also dann nur aus einem Vordersatz, wie z.B. die meisten Kinderabzählreime. Diese Zweiteilung der Periode wollen wir uns graphisch veranschaulichen.



Vordersatz und Nachsatz mit jeweils vier Takten bilden den Standardfall einer 8-taktigen Periode, wie wir sie in den Beispielen 1 und 2 vorfinden. In Beispiel 3 haben wir den Fall eines erweiterten Vordersatzes, in Beispiel 4 sind sowohl der Vorder- als auch der Nachsatz erweitert. Damit uns deutlich wird, wodurch diese Erweiterungen entstanden sind, müssen wir die Vorder- und Nachsätze in kleinere Einheiten zerlegen und nach Beziehungen zwischen diesen kleineren Einheiten suchen.
Wir hören uns jetzt noch einmal das erste Beispiel an und Ihr versucht bitte, den Vordersatz und den Nachsatz in jeweils zwei kleinere Einheiten zu zergliedern und die Grenze zwischen den Einheiten zu markieren, wie Ihr das eben schon für die Grenze zwischen Vorder- und Nachsatz gemacht habt.

Wo liegt die Grenze zwischen den beiden Einheiten, die den Vordersatz und den Nachsatz in zwei Teile zergliedern?

Wir hören uns dieses Lied gleich noch einmal an und achten jetzt einmal auf die Beziehung zwischen den beiden Teilen des Vordersatzes. Diese beiden Teile haben nämlich eine gewisse Ähnlichkeit, und Ihr versucht einmal, diese Ähnlichkeit herauszuhören und zu beschreiben.

Welche Beziehung besteht nun zwischen den beiden Teilen des Vordersatzes in dem gerade gehörten Lied?
Antwort:

Die Beziehung, die hier zwischen den beiden Hälften des Vordersatzes besteht, nennt man eine Sequenz. Eine Sequenz ist allgemein formuliert eine Wiederholung eines Melodieteils auf einer anderen Tonstufe der Tonleiter. Die Beziehung der Sequenz kann nicht nur zwischen den beiden Teilen eines Vordersatzes auftreten, sondern auch zwischen Teilen, die im Nachsatz liegen, doch dazu werden wir später noch Genaueres erfahren.

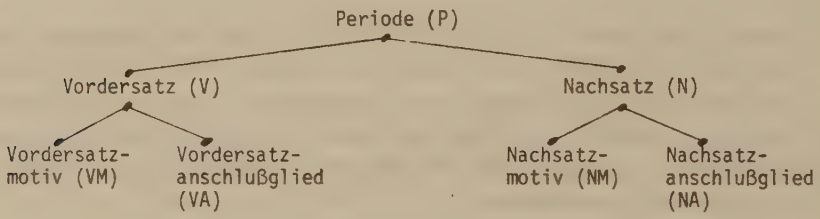
Die gerade differenzierten Teile innerhalb des Vordersatzes nennen wir Motiv des Vordersatzes und Anschlußglied des Vordersatzes oder auch Vordersatzmotiv und Vordersatzanschlußglied. Analoges gilt auch für den Nachsatz: auch hier wird die erste Hälfte als Motiv des Nachsatzes bzw. als Nachsatzmotiv bezeichnet und die zweite Hälfte als Anschlußglied des Nachsatzes bzw. als Nachsatzanschlußglied.

Im Beispiel 1 handelt es sich um den Normalfall der Untergliederung des Vordersatzes und Nachsatzes in zwei gleich große Abschnitte aus jeweils 2 Takten. Hören wir uns jetzt noch einmal Beispiel 2 an, und Ihr trefft bitte hier dieselbe Unterscheidung zwischen Motiv und Anschlußglied beim Vorder- und beim Nachsatz, indem Ihr wieder die Grenzen durch einen Strich markiert.

Wo liegen hier die Grenzen zwischen Vordersatzmotiv und -anschlußglied und Nachsatzmotiv und -anschlußglied?

Auch dieses Beispiel hören wir jetzt noch einmal, und Ihr achtet bitte wieder auf die Beziehung zwischen Motiv und Anschlußglied im Vordersatz.
Um welche Beziehung handelt es sich?
Antwort:

Wir können jetzt unser Schema der Zergliederung einer musikalischen Periode erweitern durch die Unterteilung des Vordersatzes und des Nachsatzes in ein Motiv und ein Anschlußglied.



In den Beispielen 3 und 4 hörten wir Vorder- und Nachsätze, die aus mehr als nur 4 Takten bestanden. Wir wollen uns noch einmal Beispiel 3 anhören, und dabei wollen wir uns zunächst auf den Vordersatz konzentrieren. Wir müssen jetzt herausfinden, ob der Vordersatz aus zwei Motiven und einem Anschlußglied besteht oder aus einem Motiv und zwei Anschlußgliedern. Ihr markiert bitte wieder die Grenzen zwischen Motiven und Anschlußgliedern. Die Frage lautet also:
Was ist verdoppelt, das Motiv oder das Anschlußglied des Vordersatzes?
Welche Beziehung besteht zwischen den beiden verdoppelten Teilen?
Antwort:

Die Wiederholung eines Teils einer Melodie ist eine relativ häufig anzutreffende Form der Beziehung zwischen Teilen einer Melodie. Die Beziehung der Wiederholung kann nicht nur zwischen zwei Anschlußgliedern eines Vordersatzes bestehe sondern auch zwischen anderen Einheiten, wie wir später noch sehen werden.

Die Anschlußgliedwiederholung im Vordersatz führt also zu der Erweiterung der normalerweise 8-taktigen Periode auf eine 10-taktige Periode, denn wenn eine Anschlußgliedwiederholung im Vordersatz vorkommt, darf im Nachsatz nicht auch eine Anschlußgliedwiederholung vorkommen.

Im Beispiel 4 besteht sowohl der Vordersatz als auch der Nachsatz aus jeweils 6 Takten. Wenn die eben formulierte Regel gilt, dann darf im Beispiel 4 die Erweiterung nicht durch eine Anschlußgliedwiederholung entstanden sein. Wir wollen uns dieses Beispiel noch einmal anhören, wobei Ihr wieder die Grenzen zwischen Motiven und Anschlußgliedern im Vorder- und Nachsatz markiert. Wir wollen nun wieder entscheiden, ob die Erweiterung durch eine Verdoppelung des Motivs oder eine Verdoppelung des Anschlußglieds im Vorder- und Nachsatz entstanden ist und gleichzeitig auf die Form der Beziehung achten, die zwischen den verdoppelten Teilen besteht.

Was wurde verdoppelt, die Motive oder die Anschlußglieder?
Antwort:

Welche Beziehung besteht zwischen den beiden Motiven sowohl im Vorder- als auch im Nachsatz?
Antwort:

Bei diesem Beispiel handelt es sich also um eine Erweiterung der Periode auf 12 Takte durch Motivsequenz im Vorder- und Nachsatz.

Regel: Wenn die Periode aus mehr als 8 Takten besteht, dann liegt eine erweiterte Periode vor. Es gibt nur zwei Arten der Erweiterung: 1. die alleinige Erweiterung des Vordersatzes durch eine Anschlußgliedwiederholung und 2. die Erweiterung von Vorder- und Nachsatz durch Motivsequenz

In den Beispielen 1 und 2 haben wir ebenfalls eine Motivsequenz, aber keine Erweiterung der Periode, d.h. in diesen Beispielen fungiert das sequenzierte Motiv als Anschlußglied des Vordersatzes. Hören wir nun noch einmal diese beiden Beispiele und achten genau auf die Motivsequenz im Vordersatz.

Wir haben die Beziehung der Motivsequenz also in zwei Funktionen kennengelernt: 1. als Vertreter des Anschlußgliedes, wie in den beiden gerade gehörten Beispielen und 2. als Erweiterung des Vorder- und Nachsatzes, wobei der Motivsequenz ein eigenständiges Anschlußglied folgt, das keine systematische Beziehung zu dem Motiv besitzt. Wenn wir also eine Motivsequenz gehört haben, dann signalisiert dies noch nicht eindeutig das Ende des Vordersatzes, denn auf die Motivsequenz könnte ja noch ein eigenständiges Anschlußglied folgen, wodurch eine auf 12 Takte erweiterte Periode entstehen würde. Andersherum ausgedrückt: Wenn wir eine Motivsequenz identifiziert haben, braucht dies noch nicht zu heißen, daß eine erweiterte Periode vorliegt, denn die Motivsequenz kann ja auch das Anschlußglied vertreten. Wenn aber einer Motivsequenz noch ein eigenständiges Anschlußglied folgt, dann handelt es sich mit Sicherheit um den Fall der durch Motivsequenz im Vorder- und Nachsatz erweiterten Periode.

Ähnlich unsicher ist es beim Fall der Anschlußgliedwiederholung. Wenn wir das Anschlußglied identifiziert haben, dann heißt dies noch nicht, daß der Vordersatz beendet ist, denn dem Anschlußglied könnte ja noch die Wiederholung des Anschlußgliedes folgen, die zu einer auf 10 Takte erweiterten Periode führt. Deshalb ist es gar nicht immer so leicht, die Grenze zwischen Vordersatz und Nachsatz herauszuhören. Wir werden im Verlauf des Trainings auch noch weitere Kriterien kennenlernen, die uns diese Aufgabe erleichtern.

Fassen wir das bis hierher gesagte noch einmal zusammen:

1. Alle Volks- und Kinderliedmelodien, mit denen wir uns hier beschäftigen wollen, bestehen aus zwei großen Abschnitten, nämlich aus dem Vordersatz und dem Nachsatz. Im Normalfall bestehen sowohl der Vordersatz als auch der Nachsatz aus jeweils 4 Takten.
2. Vorder- und Nachsatz lassen sich in zwei Unterabschnitte zergliedern, nämlich in ein Motiv und ein Anschlußglied. Motive und Anschlußglieder bestehen jeweils aus 2 Takten, wenn es sich um den Normalfall der 8-taktigen Periode handelt.
3. Zwischen den Motiven und Anschlußgliedern können Beziehungen bestehen, die eine bestimmte Art der Ähnlichkeit zum Ausdruck bringen. Von diesen Beziehungen haben wir bis jetzt die Wiederholung und die Sequenz kennengelernt.

4. Es gibt zwei Möglichkeiten der Periodenerweiterung: 1. Die Wiederholung des Anschlußgliedes im Vordersatz alleine, wodurch der Vordersatz auf 6 Takte erweitert wird und die ganze Periode auf 10 Takte. 2. Die Motivsequenz sowohl im Vorder- als auch im Nachsatz. Dadurch werden Vorder- und Nachsatz jeweils um zwei Takte erweitert, so daß eine 12-taktige Periode entsteht.

Aus diesen Regeln ergeben sich vier Fälle, wie eine zweiteilige Volksliedperiode durch Motive und Anschlußglieder aufgebaut sein kann. Wir betrachten diese Fälle noch einmal, indem wir die Gliederung der Periode als Baumstruktur darstellen und auch die verschiedenen Beziehungen in die Darstellung aufnehmen.

1. Der Normalfall der Periode besteht aus einem 4-taktigen Vordersatz und einem 4-taktigen Nachsatz. Dabei kann a) der Vordersatz und der Nachsatz aus einem Motiv und einem Anschlußglied mit jeweils 2 Takten bestehen, wobei zwischen den Motiven und Anschlußgliedern keine Beziehungen bestehen. Das folgende Beispiel 5 repräsentiert diesen Fall 1a.
Die zu diesem Fall gehörige Baumstruktur steht bereits auf S. 7 (S. 506).

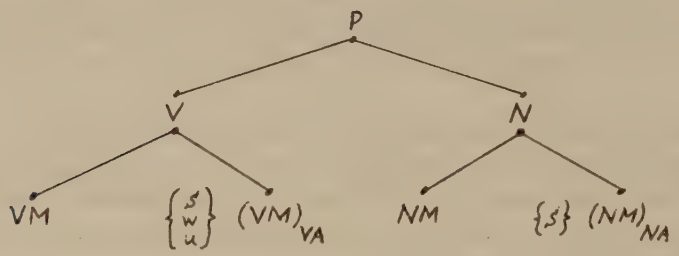
Fall 1b ist ebenfalls der Normalfall der 8-taktigen Periode, unterscheidet sich aber vom Fall 1a dadurch, daß zwischen den Motiven und Anschlußgliedern Beziehungen bestehen können, ohne daß die Periode dadurch erweitert würde. Von diesen Beziehungen haben wir bereits die Sequenz und die Wiederholung kennengelernt. Beispiel 6 ist noch einmal ein Beispiel für diesen Fall 1b mit einer Motivsequenz im Vordersatz.

Die zweite Art der Beziehung zwischen Motiv und Anschlußglied ist die der Wiederholung. Hierfür haben wir noch kein musikalisches Beispiel gehört. Da aber Wiederholungen ohnehin am einfachsten zu erkennen sind, wollen wir hier auf ein Beispiel verzichten.

Die dritte Art der Beziehung zwischen Motiv und Anschlußglied haben wir noch nicht kennengelernt. Hierbei handelt es sich um die Beziehung der Umkehrung, die nur zwischen Motiv und Anschlußglied des Vordersatzes vorkommt. Beispiel 7 enthält diese Umkehrung zwischen Motiv und Anschlußglied. Wir hören hier nur den Vordersatz.

Wie wir an diesem Beispiel gesehen haben, besteht die Umkehrung aus einer Wiederholung der Töne des Vordersatzmotivs in umgekehrter Reihenfolge als Vordersatzanschlußglied, wobei zwischen Motiv und Anschlußglied die für die Umkehrung charakteristische Symmetrie der Melodie entsteht.

Die folgende Baumstruktur für diesen Fall 1b gibt die erlaubten Beziehungen zwischen Motiven und Anschlußgliedern wieder.

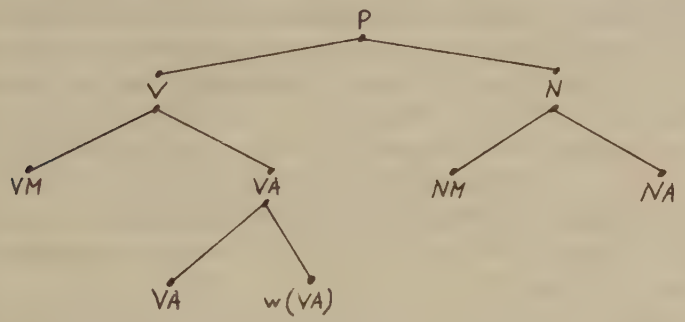


Die in der geschweiften Klammer stehenden Kleinbuchstaben stehen für die alternativen Beziehungen: s für Sequenz, w für Wiederholung und u für Umkehrung. In der runden Klammer steht nun statt VA für Vordersatzanschlußglied VM für Vordersatzmotiv, und als Index steht VA. Dies ist nun folgendermaßen zu lesen: An der Stelle, wo das Vordersatzanschlußglied stehen muß, steht hier die Tonfolge des Vordersatzmotivs in der abgewandelten Form, nämlich entweder als Sequenz des Motivs, als Wiederholung oder als Umkehrung. Das Analoge gilt für den Nachsatz. Diese Baumstruktur enthält jetzt alle möglichen Beziehungen zwischen Motiven und Anschlußgliedern, die bei unseren Melodien vorkommen können.

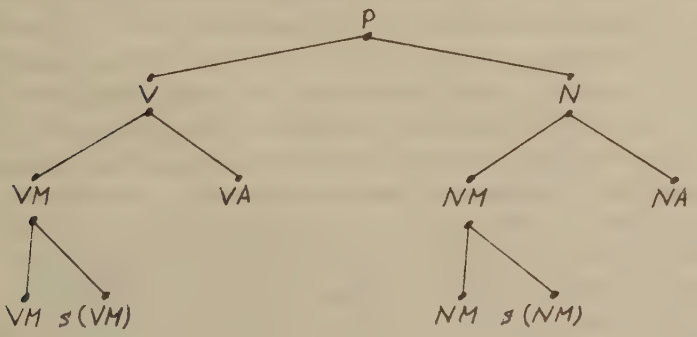
Regel: Zwischen Vordersatzmotiv und Vordersatzanschlußglied können die drei Beziehungen Sequenz, Wiederholung und Umkehrung vorkommen, zwischen Nachsatzmotiv und Nachsatzanschlußglied aber nur die Beziehung der Sequenz.

Ich spiele Euch jetzt drei Beispiele vor, bei denen Ihr bitte herauszufinden versucht, ob das Vordersatzanschlußglied eine Umkehrung des Vordersatzmotivs ist oder nicht. Bitte schreibt neben die Nummer der Beispiele 8 bis 10 ein kleines u, wenn Ihr eine Umkehrung zu erkennen meint.

Wir gehen nun über zu den Fällen, in denen die Periode auf 10 oder 12 Takte erweitert wurde. Fall 2a ist der Fall, in dem der Vordersatz alleine erweitert wurde. Dies geschieht ausschließlich durch eine Wiederholung des Vordersatzanschlußgliedes. Die Struktur dieser 10-taktigen Periode gibt der folgende Strukturbaum wieder.



Fall 2b ist der Fall, in dem Vordersatz und Nachsatz durch eine Motivsequenz erweitert werden, so daß eine 12-taktige Periode entsteht. Die Struktur der 12-taktigen Periode gibt der folgende Baum wieder.



Wir hören uns jetzt vier Beispiele an, und Ihr entscheidet bitte in jedem Fall, ob es sich um eine Erweiterung des Vordersatzes alleine durch eine Anschlußgliedwiederholung handelt oder um den Fall einer Erweiterung von Vorder- und Nachsatz durch eine Motivsequenz. Damit Ihr es nicht zu einfach habt und diese Entscheidung schon alleine aufgrund der Zahl der Takte fällen könnt, spiele ich nur die Vordersätze. Auf S. 2 des Arbeitshefts, Beispiel 11 bis 14, findet Ihr die Noten. Macht zur Beantwortung ein Kreuz bei w(VA), wenn es sich um eine Erweiterung aufgrund einer Anschlußgliedwiederholung handelt, und bei s(VM), wenn es sich um eine Erweiterung aufgrund einer Motivsequenz handelt.

Bevor wir das neuerworbene Wissen etwas intensiver anwenden und uns im Erkennen dieser verschiedenen Strukturen üben, wollen wir eine viertelstündige Pause einlegen.

(Pause von 15 Minuten)

Die Struktur der Lieder zu erkennen, ist natürlich einfacher, wenn man die Noten vor sich hat. Wir wollen jetzt aber das Heraushören der Strukturen trainieren, wozu Euch 14 Lieder vorgespielt werden. Eure Aufgabe ist es, auf dem Antwortblatt auf S. 3 des Arbeitshefts anzukreuzen, ob es sich bei den einzelnen Melodien 1. um den Normalfall einer 8-taktigen Periode handelt oder nicht, und wenn ja, sollt Ihr zusätzlich entscheiden, ob zwischen dem Vordersatzmotiv und dem -anschlußglied eine Beziehung besteht oder nicht, und um welche Form von Beziehung es sich jeweils handelt. Wir haben die Beziehung der Wiederholung fortgelassen, weil diese ohnehin einfach zu erkennen ist und wir uns deshalb diesen Fall sparen können. Wenn es sich bei einer Melodie um den 2. Fall einer erweiterten Periode handelt, dann müßt Ihr entscheiden, ob es sich um den Fall der Erweiterung durch Motivsequenz oder um den Fall der Erweiterung durch Anschlußgliedwiederholung handelt.

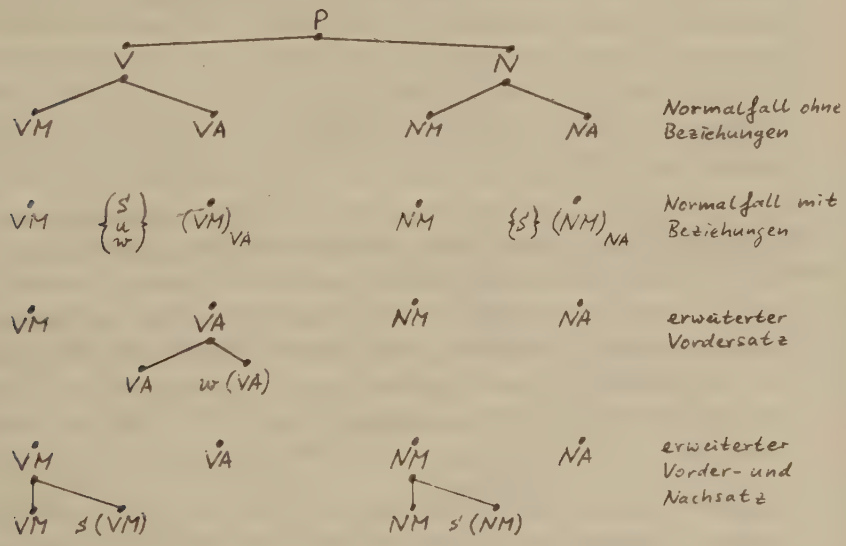
Um die Entscheidung zwischen den beiden Fällen der Erweiterung nicht zu leicht zu machen, werden immer nur 6 Takte vorgespielt und dann ausgeblendet.

Zweite Trainingssitzung

Phrasenstrukturen unterhalb des Subphrasenniveaus

Wir haben in der ersten Sitzung bereits einen großen Teil der Strukturen kennengelernt, die in zweiteiligen Volks- und Kinderliedern anzutreffen sind. Zunächst wollen wir kurz rekapitulieren, welche Teilstrukturen wir bereits kennengelernt haben.

Die einzelnen Schritte der Zergliederung einer Volksliedmelodie können wir uns am besten anhand eines Strukturbaums veranschaulichen, wie wir ihn bereits in der letzten Sitzung mehrfach verwendet haben.



Die gesamte Melodie haben wir als musikalische Periode bezeichnet und durch den Buchstaben P symbolisiert. Die Periode teilt sich in zwei musikalische Phrasen, nämlich in den Vordersatz und den Nachsatz (Symbole V und N). Beide Phrasen V und N bestehen ihrerseits aus wenigstens zwei Subphrasen, nämlich jeweils aus einem Motiv und einem Anschlußglied. Die Subphrasen, also die Motive und Anschlußglieder, bestehen immer aus zwei Takten.

Aufgrund der Anzahl der in einer Periode vorkommenden Subphrasen und der Beziehungen, die zwischen ihnen bestehen, haben wir vier verschiedene Fälle unterschieden:

1a) Normalfall der 8-taktigen Periode, bestehend aus Vordersatz und Nachsatz jeweils mit einem Motiv und einem Anschlußglied ohne Beziehungen zwischen Motiven und Anschlußgliedern. Ich spiele Euch hierfür ein Beispiel vor, das Ihr auf S. 4 des Arbeitshefts als Beispiel 15 findet.

1b) Beim zweiten Fall handelt es sich ebenfalls um den Normalfall der 8-taktigen Periode, aber bei diesem Fall bestehen zwischen den Motiven und Anschlußgliedern Beziehungen. Beim Vordersatz kann es sich um die Beziehung der Sequenz (s), der Umkehrung (u) oder der Wiederholung (w) handeln, beim Nachsatz nur um die Beziehung der Sequenz. Wir hören nun Beispiel Nr. 16 mit einer Motivsequenz im Vordersatz und gleich anschließend Beispiel Nr. 17 mit einer Motivumkehrung im Vordersatz. (Noten auf S. 4 des Arbeitshefts!)

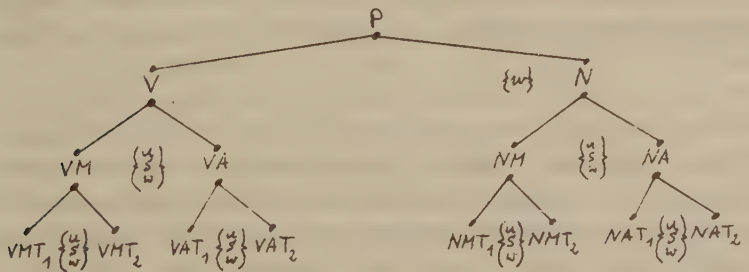
Beim dritten und vierten Fall handelt es sich um die beiden Möglichkeiten der Erweiterung einer Periode. Fall 2a ist der Fall der Erweiterung der Periode von 8 auf 10 Takte durch Anschlußgliedwiederholung im Vordersatz alleine. Fall 2b ist der Fall der Erweiterung der Periode auf 12 Takte durch Motivsequenz im Vorder- und Nachsatz. Beim folgenden Beispiel 18 handelt es sich um den Fall der Erweiterung durch Anschlußgliedwiederholung und beim Beispiel 19 um den Fall der Erweiterung durch Motivsequenz. (Noten auf S. 4 des Arbeitshefts!)

Zwischen Vordersatz und Nachsatz haben wir bisher noch keine Beziehung kennengelernt. Es gibt aber eine Beziehung, die wir jetzt noch nachtragen wollen. Dabei handelt es sich um den Fall der Wiederholung des Vordersatzes als Nachsatz mit Abwandlung am Schluß. Beispiel 20 enthält diese Beziehung (S. 4 des Arbeitshefts!).

Diese Wiederholung des Vordersatzes mit Abwandlung am Schluß gibt es nur im Normalfall der 8-taktigen Periode. Meistens werden die ersten drei Takte des Vordersatzes identisch für die ersten drei Takte des Nachsatzes übernommen, so daß die Abwandlung am Schluß sich nur auf den letzten Takt bezieht. Wenn zwischen Vordersatz und Nachsatz die Beziehung der Wiederholung mit Abwandlung am Schluß besteht, können zusätzlich aber auch noch Beziehungen zwischen den Motiven und Anschlußgliedern bestehen. Dies soll uns das Beispiel 21 verdeutlichen. Welche Beziehung besteht dort zusätzlich zur Beziehung der Wiederholung mit Abwandlung am Schluß? (Noten auf S. 4 des Arbeitshefts!)

Antwort:

Wir wollen die Gliederung der Melodien nun weiter nach unten fortsetzen, d.h. wir wollen jetzt musikalische Einheiten betrachten, die kleiner als die Subphrasen sind. Da Motive und Anschlußglieder immer aus zwei Takten bestehen, sind die nächst kleineren Einheiten die uns bereits bekannten Takte. Zwischen Takten, die zu einem Motiv oder einem Anschlußglied gehören, können all die Beziehungen bestehen, die wir bereits für die größeren Einheiten kennengelernt haben, nämlich die Beziehung der Wiederholung, die der Sequenz und die der Umkehrung. Der folgende Strukturbaum für den Fall der 8-taktigen Periode veranschaulicht dies.



Ich spiele Euch jetzt einige Beispiele vor, die aus den beiden Takten des Vordersatzmotivs bestehen. Diese Beispiele findet Ihr auf S. 5 Eures Arbeitshefts in Notenschrift. Beispiel 22 ist ein Fall der Wiederholung.

Das folgende Beispiel ist ein Fall der Sequenz, und zwar eine Sequenz um eine Sekunde aufwärts.

Beispiel 24 stellt ebenfalls eine Sequenz dar, aber jetzt eine Quarte abwärts.

Die Beispiele 25 und 26 enthalten jeweils eine Umkehrung.

Diese verschiedenen Beziehungen können an verschiedenen Stellen frei kombiniert auftreten, d.h. z.B. kann zwischen den beiden Takten des Vordersatzmotivs die Beziehung der Sequenz bestehen und gleichzeitig zwischen den beiden Takten des Anschlußglieds im Vordersatz die Beziehung der Umkehrung, und im Nachsatz kann zwischen den beiden Takten des Motivs z.B. eine Wiederholung bestehen.

Regel: Zwischen den Takten der Motive und Anschlußglieder können die Beziehungen der Sequenz, Umkehrung oder Wiederholung bestehen.

Auch die Takte können wir noch weiter unterteilen, und zwar nennt man diese Elemente die Füße eines Taktes. Wir beschränken uns hier ganz auf den 4/4 - Takt.

Im 4/4 - Takt gibt es zwei Füße, nämlich den 1., der die Hauptbetonung auf der 1. Zählzeit (auf dem 1. Viertel) trägt und dem eine unbetonte Zählzeit folgt, und den 2. Fuß, der eine Nebenbetonung (zweitstärkste Betonung) auf der ersten Zählzeit des Fußes (= 3. Zählzeit des Takts) trägt und dem ebenfalls eine unbetonte Zählzeit folgt. Die 1. Zählzeit eines Taktes trägt also die Hauptbetonung, der eine unbetonte Zählzeit folgt, und diese beiden Zählzeiten bilden zusammen den ersten Fuß. Die dritte Zählzeit eines Taktes trägt die Nebenbetonung, der eine unbetonte Zählzeit folgt, und diese beiden bilden zusammen den zweiten Fuß.

Im folgenden Beispiel eines Liedes haben wir die Betonungen einmal überdeutlich gespielt, damit Ihr deutlich die beiden Füße eines Taktes heraushören könnt. Die Symbole über den Noten geben die relative Stärke der Betonungen an. (Beispiel 23 auf S. 5 des Arbeitshefts!)

Auch zwischen den Füßen eines Taktes können die uns bereits bekannten Beziehungen vorkommen, nämlich die Wiederholung (Beispiel 28), die Sequenz (Beispiel 29), die Umkehrung (Beispiel 30) und natürlich kann es auch den Fall keiner Beziehung geben (Beispiel 31). Diese Beispiele hören wir uns jetzt hintereinander an, und Ihr vergleicht bitte mit dem Notenbild.

Beispiel 32 stellt einen sehr häufigen Fall von Beziehung zwischen Füßen dar, nämlich eine in gleich großen Intervallen absteigende (oder aufsteigende) Melodielinie. Um welche Art von Beziehung zwischen Füßen eines Taktes handelt es sich hierbei?

Antwort:

Nun wollen wir das Erkennen der Beziehungen zwischen Takten und Füßen ein wenig üben. Ich spiele Euch jetzt 14 zweitaktige Melodieausschnitte vor, und Ihr entscheidet bitte, auf welchem Niveau, d.h. ob zwischen Takten oder Füßen, eine Beziehung besteht und um welche Beziehung es sich jeweils handelt. Achtet bitte darauf, daß sowohl auf Taktniveau als auch auf Fußniveau gleichzeitig Beziehungen bestehen können. Für Eure Antwort verwendet bitte das Antwortblatt auf S. 6 des Arbeitshefts. Macht bitte ein Kreuz in das entsprechende Kästchen, wenn nur auf einem Niveau eine Beziehung vorliegt, und zwei Kreuzchen in die entsprechenden Kästchen, wenn auf zwei Niveaus Beziehungen vorliegen.

(15 Minuten Pause!)

Bevor wir jetzt das Erkennen und Unterscheiden von Beziehungen auf allen Niveaus einer musikalischen Periode üben, wollen wir noch einmal ganz kurz die einzelnen Begriffe wiederholen, die die Elemente einer Melodie auf den verschiedenen Niveaus bezeichnen, denn diese Begriffe braucht Ihr gleich, um die Fragen richtig zu verstehen, die zu jeder vorgespielten Melodie gestellt werden.

Das alleroberste Niveau ist die Periode, die aus 8, 10 oder 12 Takten bestehen kann. Darunter liegt das Niveau der Phrasen. Jede Periode bei den von uns verwendeten Melodien besteht aus zwei Phrasen, nämlich dem Vordersatz und dem Nachsatz. Jede Phrase teilt sich auf dem darunter liegenden Niveau in Subphrasen auf, nämlich in wenigstens ein Motiv und ein Anschlußglied. Es gibt also jeweils ein Vordersatzmotiv und ein Vordersatzanschlußglied sowie ein Nachsatzmotiv und ein Nachsatzanschlußglied. Vergleicht bitte mit dem Strukturbaum auf S. 16. Motive und Anschlußglieder werden auf dem darunter liegenden Niveau in die beiden Takte eines Motivs bzw. eines Anschlußglieds aufgeteilt. Jeder Takt besteht auf dem niedrigsten Niveau aus zwei Füßen.

Wenn jetzt nach der Beziehung zwischen den beiden Takten des Vordersatzmotivs gefragt wird, dann sind die beiden ersten Takte gemeint. Wenn eine Frage nach der Beziehung zwischen den beiden Füßen des 2. Takts des Nachsatzanschlußglieds gestellt wird, dann sind die Füße des letzten Takts gemeint.

Ich spiele Euch jetzt 15 Melodien vor, wobei bei einigen Melodien nach dem 6. Takt ausgeblendet wird. Zu jeder Melodie habt Ihr auf S. 7 Eures Arbeitshefts eine Frage nach der Beziehung auf einem ganz bestimmten Niveau. Nur das in der Frage bezeichnete Niveau soll die Grundlage Eurer Entscheidung sein, obwohl gleichzeitig auch noch auf anderen Niveaus Beziehungen bestehen können. Antwortet auf jede Frage bitte durch Ankreuzen oder Unterstreichen der richtigen Antwortalternative. Lest jetzt bitte die erste Frage auf S. 7 !

Dritte Trainingssitzung

Harmonische und melodische Regelmäßigkeiten

Wir wollen uns heute mit einigen harmonischen und melodischen Merkmalen von Volksliedmelodien beschäftigen, die es uns gestatten, auch dann Einheiten auf verschiedenen Niveaus der Melodiestruktur voneinander abzugrenzen, wenn keine Beziehungen in Form von Wiederholungen, Sequenzen oder Umkehrungen zwischen ihnen festzustellen sind. Dazu müssen wir ein wenig Harmonielehre wiederholen, wovon Euch das meiste bereits vertraut sein dürfte.

Für jede Tonart gibt es einen Hauptbezugspunkt für alle Töne, und das ist der Grundton der Tonleiter. Der Grundton einer Tonleiter ist der 1. Ton der Tonleiter, der sich bei aufsteigend gespielter Tonleiter eine Oktave höher wiederholt. Ich spiele Euch zuerst noch einmal die C-Dur-Tonleiter vor und gleich anschließend eine Melodie, in der der Grundton mehrfach vorkommt. Bitte markiert auf S. 10 Eures Arbeitshefts in Beispiel 33 den Grundton mit einem Strich oder Kreuzchen, sobald er auftritt.

Wenn am Ende des Nachsatzes der Grundton auftritt, wie dies meistens der Fall ist, dann löst er die im Verlauf der Melodie erzeugte Spannung auf und signalisiert so besonders deutlich das Ende einer Periode. Wegen dieser Funktion der Spannungsauflösung wird der Grundton zum beherrschenden Ton einer ganzen Melodie: am Anfang strebt die Melodie vom Grundton fort und bewegt sich am Ende wieder auf den Grundton zu. Typische Volks- und Kinderlieder beginnen auf dem Grundton und enden auf dem Grundton. Dies ist aber leider nicht immer so, so daß es manchmal nicht ganz einfach ist, das Ende einer Periode festzustellen bzw. zu entscheiden, ob die Melodie einfach abgebrochen wurde oder tatsächlich zu Ende ist. Wenn die Periode nicht auf dem Grundton beginnt oder endet, so beginnt oder endet sie doch mit einem Ton, der ein besonderes Verwandtschaftsverhältnis zum Grundton besitzt. Dieses Verwandtschaftsverhältnis ist die Zugehörigkeit eines Tones zu einer bestimmten Harmonie. Wir müssen also jetzt feststellen, zu welcher Harmonie die Töne gehören, mit denen eine Periode anfangen und enden kann.

Alle Töne einer Tonleiter werden in zweifacher Hinsicht auf den Grundton bezogen: 1. bezüglich ihrer Höhe über oder unter dem Grundton, d.h. die Tonstufen werden relativ zum Grundton angegeben. Die dritte Tonstufe innerhalb einer Tonleiter ist also der Ton, der an dritter Stelle einer aufsteigenden Tonleiter steht. 2. Die zweite Beziehung zum Grundton ist die des harmonischen Zusammenklangs mit dem Grundton oder mit dem 5. Ton über dem Grundton oder dem 5. Ton unter dem Grundton.

Die Töne, die mit dem Grundton einen harmonischen Zusammenklang bilden, werden die Töne des Tonikadreiklangs genannt. Dies sind die Töne der 3. und 5. Tonstufe, in C-Dur also das e und das g. Beispiel 34 auf S. 10 des Arbeitshefts enthält einige Dreiklänge der Tonika, einmal simultan als Akkord gespielt und dann sukzessiv als Tonfolge gespielt.

Die Töne, die mit der 5. Tonstufe über dem Grundton harmonisch zusammenklingen, sind die Töne der 2. und 7. Tonstufe, also d und h in C-Dur. Sie werden die Töne des Dominantakkords genannt. Beispiel 35 enthält solche Dominantakkorde bzw. Tonfolgen mit den Tönen des Dominantakkords.

Die Töne, die mit der 5. Tonstufe unterhalb des Grundtons (= 4. Tonstufe über dem Grundton) harmonisch zusammenklingen, sind die Töne der 1. und der 6. Tonstufe, also c und a in C-Dur. Sie werden die Töne des Subdominantakkords genannt. Beispiel 36 enthält Subdominantakkorde bzw. Tonfolgen der Töne des Subdominantakkords.

Wir wollen nun einmal versuchen, die Töne in einer Melodie nach ihrer harmonischen Zugehörigkeit zu kategorisieren. Im folgenden Beispiel 37 besteht jeder Takt aus Tönen, die alle zur gleichen Harmonie gehören, also entweder alle zur Tonika, oder alle zur Subdominante oder alle zur Dominante. Bitte schreibt beim Hören der Melodie unter jeden Takt, zu welcher Harmonie die Töne des Takts gehören. Wenn es sich um die Töne der Tonika (c, e, g) handelt, dann schreibt ein großes T unter den Takt, wenn es sich um die Töne der Subdominante (f, a, c) handelt, dann schreibt ein großes S unter den Takt, und wenn es sich um die Töne der Dominante handelt (g, h, d), dann schreibt ein großes D unter den Takt.

In echten Volksliedern ist es leider selten so einfach, die harmonische Zugehörigkeit einzelner Töne zu bestimmen, wie in unserem letzten Beispiel. Häufig ist es so, daß nur die beiden Töne eines Fußes zur gleichen Harmonie gehören und die Töne des nächsten Fußes zu einer anderen Harmonie, d.h. es tritt von Fuß zu Fuß ein Harmoniewechsel auf. Im folgenden Beispiel 41 auf S. 11 des Arbeitshefts findet bei fast jedem Fuß ein Harmoniewechsel statt.

Wenn die Töne eines Fußes aber immer beide Harmonietöne sein müßten, wäre das Intervall zwischen dem Ton auf der betonten Zählzeit und der unbetonten Zählzeit entweder immer eine Terz oder eine Quinte, was zu ziemlich langweiligen Melodien führen würde. In Volks- und Kinderliedern sind auch sogenannte harmoniefremde Töne erlaubt, von denen wir hier nur zwei Typen erwähnen wollen. Harmoniefremde Töne sind Töne auf unbetonter Zählzeit, die nicht zur Harmonie des auf betonter Zählzeit stehenden Tons gehören. Das folgende Melodiebeispiel 42 enthält auf den meisten unbetonten Zählzeiten sogenannte Durchgangstöne.

Durchgangstöne sind also harmoniefremde Töne auf unbetonter Zählzeit, die im Sekundabstand zwischen zwei auf betonter Zählzeit stehenden Harmonietönen stehen. Der zweite Typ von harmoniefremden Tönen ist der der Wechseltöne. Das folgende Beispiel enthält auf mehreren unbetonten Zählzeiten solche Wechseltöne.

Soviel zur Wiederholung einiger Grundlagen der Harmonielehre. Wir wollen jetzt diese Begriffe zur Charakterisierung einiger harmonischer Regelmäßigkeiten in Volksliedmelodien heranziehen. Wir fangen am besten beim Anfang des Vordersatzes an und schauen uns noch einmal einige der bereits gehörten Beispiele an. Bitte sucht einmal in beliebigen Beispielen nach den Harmonien, die beim ersten Fuß des Vordersatzes vorkommen. (ca. 5 Min. Zeit!)

Welche Harmonien habt Ihr für den ersten Fuß des Vordersatzes gefunden?
Antwort:

Jetzt machen wir das gleiche für den letzten Fuß des Vordersatzes, welche Harmonien kommen am Ende des Vordersatzes vor? (ca. 5 Min. Zeit!)
Antwort:

Also fassen wir zusammen: der Vordersatz beginnt grundsätzlich beim 1. Fuß nur mit Tönen der Tonika und endet beim letzten Fuß grundsätzlich nur mit Tönen der Tonika oder der Dominante, wobei von den Tönen der Dominante nur die der 2. und 5. Tonstufe vorkommen. Zwischendurch sind alle Harmoniefolgen erlaubt, allerdings mit einer Einschränkung: die Harmoniefolge Dominante - Subdominante ist nicht erlaubt.
Jetzt schauen wir uns einmal die Anfänge der Nachsätze an, welche Harmonien kommen dort vor?

Antwort:

Und welche Harmonien kommen am Ende des Nachsatzes vor? (ca. 5 Min. Zeit!)
Antwort:

Fassen wir noch einmal zusammen: Die musikalische Periode eines zweiteiligen Volks- oder Kinderliedes verläuft harmonisch von der Tonika im Vordersatzanfang über beliebige Harmonien (mit Ausnahme von Dominant-Subdominant-Folgen) zu einem vorläufigen Ruhepunkt am Ende des Vordersatzes, der durch die Tonika oder die Dominante gebildet werden kann. Der Nachsatz beginnt mit einer beliebigen Harmonie und endet ausnahmslos auf einem Ton der Tonika, meistens auf dem Grundton.

Beim Vordersatzende und beim Nachsatzende haben wir immer von einem Ruhepunkt gesprochen. Wenn wir uns die bereits gehörten Beispiele ansehen, fällt auf, daß entweder an diesen Stellen ein doppelt langer Notenwert steht oder aber derselbe Ton wiederholt wird. Auch an dieser rhythmischen bzw. melodischen Besonderheit ist das Ende des Vorder- und Nachsatzes recht gut zu erkennen.

Die Grenze zwischen Motiven und Anschlußgliedern ist harmonisch nicht besonders ausgezeichnet. Aber der Melodieverlauf zeigt zwischen Motiv und Anschlußglied häufig einen typisch kontrastierenden Verlauf, d.h. wenn z.B. das Vordersatzmotiv überwiegend aufsteigend verläuft, dann verläuft das Anschlußglied überwiegend absteigend oder umgekehrt. Um einen Eindruck von solchen typischen Verläufen zu bekommen, die übrigens auch zwischen Takten bestehen können, hören wir uns jetzt 10 Beispiele an. Eure Aufgabe dabei ist es, den Verlauf der Melodie in den einzelnen Takten des Vorder- und Nachsatzes zu charakterisieren, indem Ihr in die vorgefertigten Kästchen auf S. 12 des Arbeitshefts einen aufsteigenden Strich zeichnet, wenn die Melodie in diesem Takt überwiegend aufsteigend verläuft bzw. einen absteigenden Strich zeichnet, wenn die Melodie überwiegend absteigend verläuft, sollten auf- und absteigende Teile einander die Waage halten oder die Tonhöhe immer gleich bleiben, dann macht einfach einen horizontalen Strich. Beim ersten Beispiel macht der Versuchsleiter Euch die Aufgabe an der Tafel vor.

Achtet bitte besonders auf das Ende der Vordersätze, und gebt bitte bei jeder Melodie an, ob der Vordersatz auf der Tonika oder der Dominante endet, indem Ihr das richtige Harmoniesymbol unterstreicht. Ihr hört jede Melodie zweimal, damit Ihr diese nicht ganz einfache Aufgabe gut lösen könnt.
Wir machen jetzt beim 2. Lied weiter!

(Pause von 15 Min.!)

Zum Abschluß der heutigen Sitzung wollen wir das gesamte bis heute angesammelte Wissen über Volksliedstrukturen noch einmal anwenden. Auf den Seiten 13 und 14 des Arbeitshefts findet Ihr Fragen, die Ihr bitte jeweils vor dem Anhören einer Melodie aufmerksam lest. Eure Antwort gebt Ihr durch Unterstreichen oder Ankreuzen der richtigen Antwortalternative.

Vierte Trainingssitzung

Abschlußtest

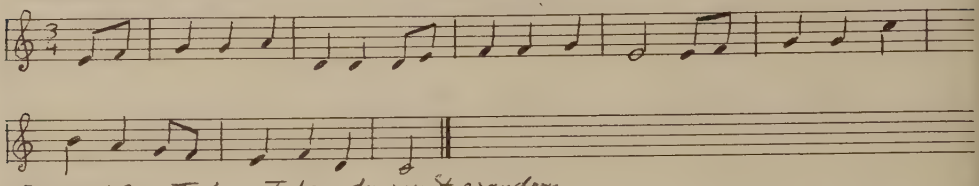
Mit der heutigen Sitzung beschließen wir das Training. Ihr habt bis heute die wichtigsten Regelmäßigkeiten in zweiteiligen Volks- und Kinderliedern kennengelernt, die Ihr teilweise in den später folgenden Versuchen anwenden müßt. Die heutige Sitzung dient dazu, etwas intensiver als bisher die Anwendung des erworbenen Wissens zu üben.

Auf S. 15 des Arbeitshefts findet Ihr die Fragen und Antwortalternativen, die Ihr Euch bitte jeweils vor dem Anhören einer Melodie aufmerksam durchlest. Am Ende einer Melodie beantwortet bitte so schnell wie möglich die gestellte Frage durch Ankreuzen oder Unterstreichen der richtigen Antwortalternative. Im Unterschied zu den vorangegangenen Trainingssitzungen gibt der Versuchsleiter im Anschluß an eine Melodie keine Rückmeldung über die richtige Antwort. Nach der Hälfte der Melodien legen wir eine viertelstündige Pause ein.

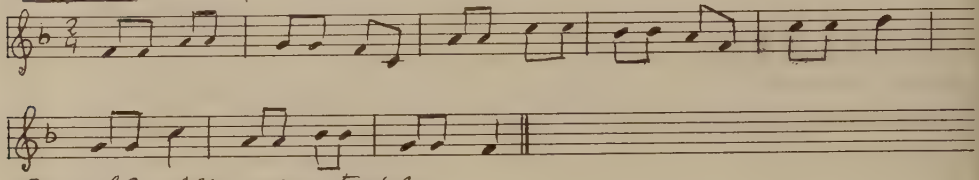
A R B E I T S H E F T

zum S y n t a x t r a i n i n g

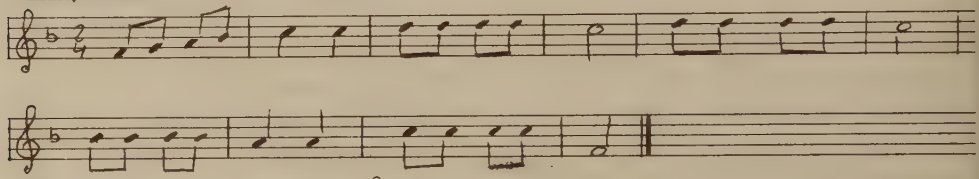
Beispiel 1: Schneeflöckchen, Weißröckchen (bei allen Beispielen ♩ = 120)



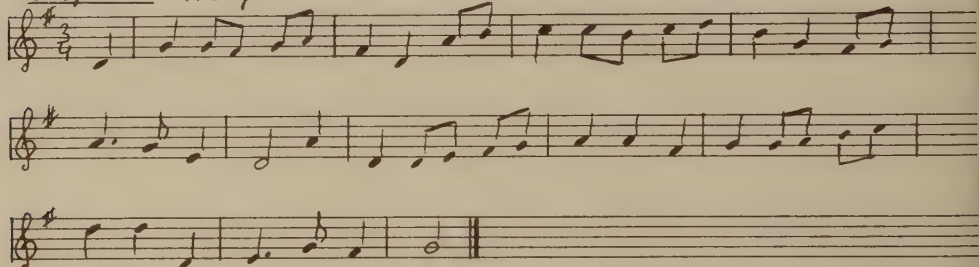
Beispiel 2: Täler, Täler, du mußt wandern



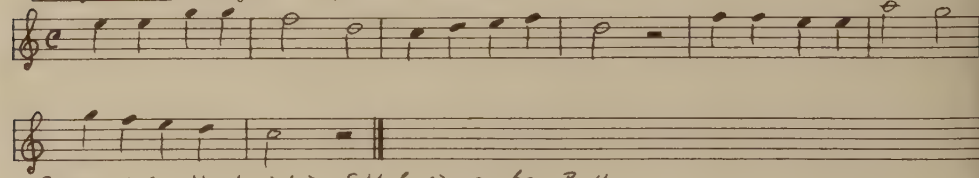
Beispiel 3: Alle meine Entchen



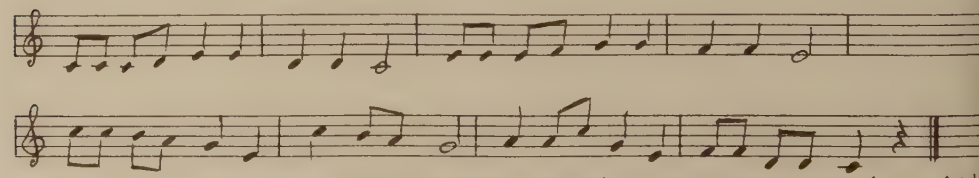
Beispiel 4: Die güldene Sonne



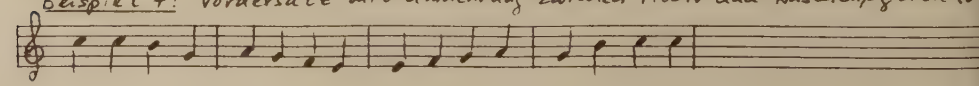
Beispiel 5: Auf dem grünen Rasen

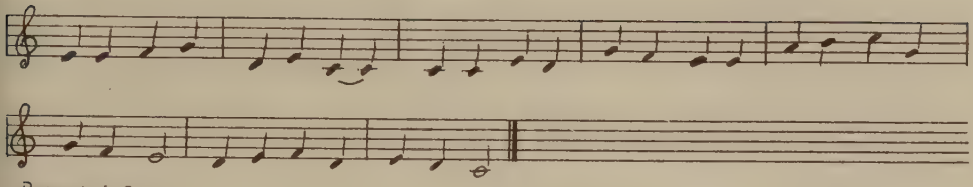


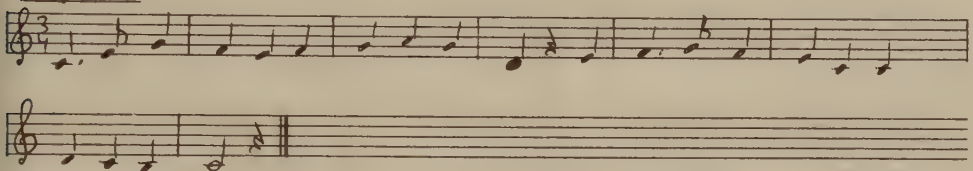
Beispiel 6: Heute ist im Schloß ein großer Ball

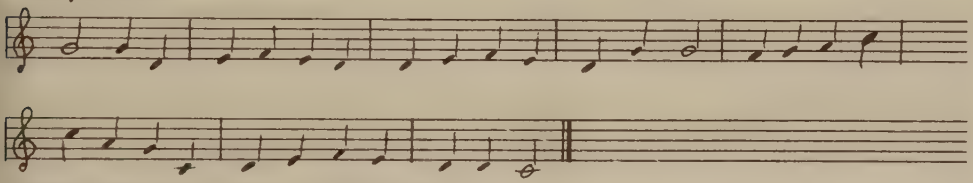


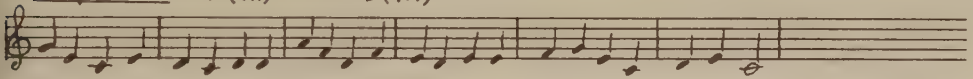
Beispiel 7: Vordersatz mit Umkehrung zwischen Motiv und Anschlußglied (♩=



Beispiel 8:

Beispiel 9:



Beispiel 10:

Beispiel 11: W (VA) oder S (VH)



Beispiel 12: W (VA) oder S (VH)

Beispiel 13: W (VA) oder S (VH)

Beispiel 14: W (VA) oder S (VH)

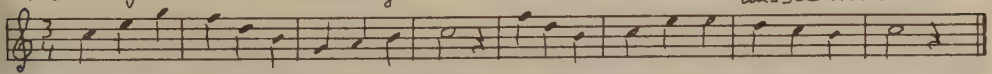
Antwortblatt für Hörübung am Ende der ersten Sitzung

A n t w o r t k a t e g o r i e n

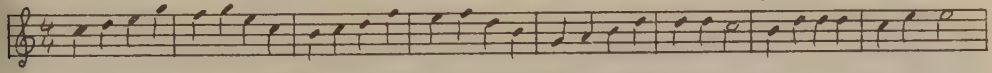
	NORMALTEIL: Beziehung zwischen Motiv u. Anschlußglied im Vordersatz			ERWEITERTE PERIODE: Art der Erweiterung	
	Keine	Übehrung	Sequenz	Motivsequenz	Anschlußgliedwiederholung
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					

Melodien für Hörübung am Ende der ersten Sitzung

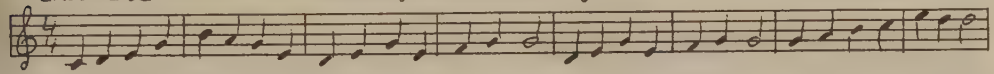
Normalfall ohne Beziehung — ausblenden!



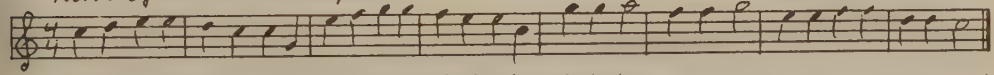
erweiterte Periode mit Motivsequenz — ausblenden!

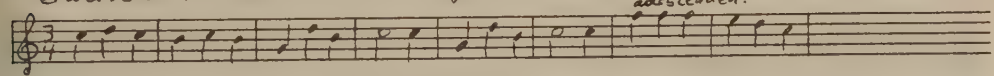


erweiterte Periode mit Anschlußgliedwiederholung — ausblenden!

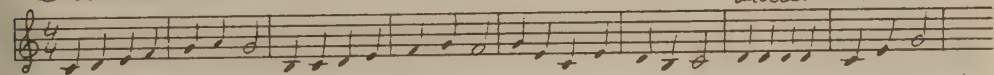


Normalfall mit Motivsequenz — ausblenden!

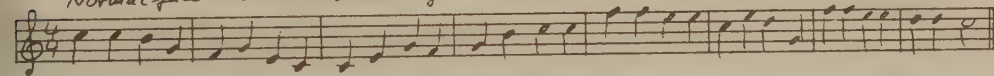


erweiterte Periode mit Anschlußgliedwiederholung — ausblenden!

erweiterte Periode mit Motivsequenz — ausblenden!

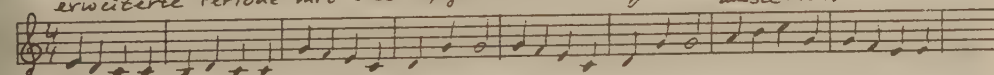
erweiterte Periode mit Motivsequenz — ausblenden!



Normalfall mit Motivumkehrung — ausblenden!

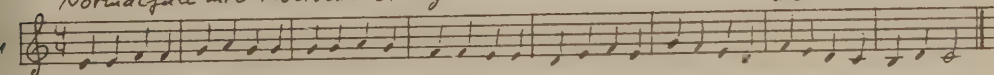


erweiterte Periode mit Anschlußgliedwiederholung — ausblenden!

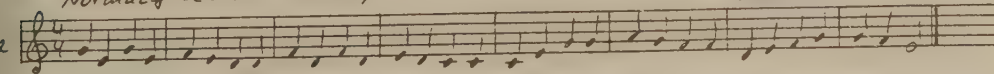


erweiterte Periode mit Motivsequenz — ausblenden!

10

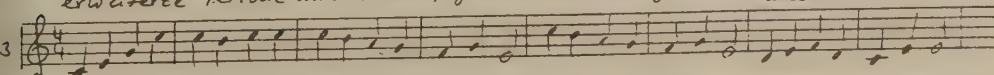
Normalfall mit Motivumkehrung — ausblenden!

11

Normalfall mit Motivsequenz — ausblenden!



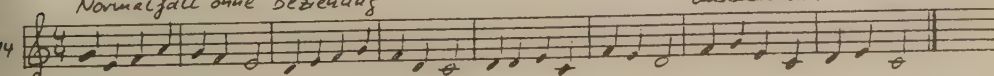
12

erweiterte Periode mit Anschlußgliedwiederholung — ausblenden!



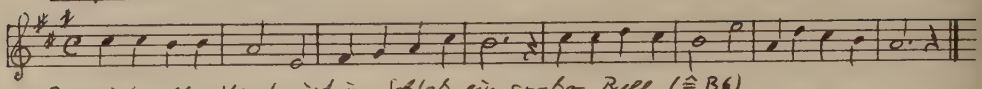
13

Normalfall ohne Beziehung — ausblenden!

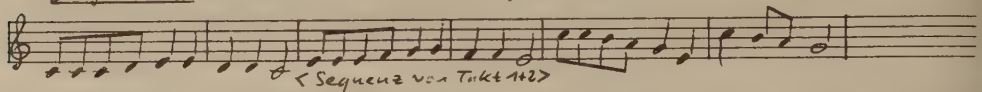


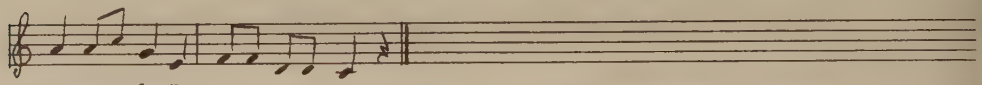
14

Beispiel 15: Abend wird es wieder

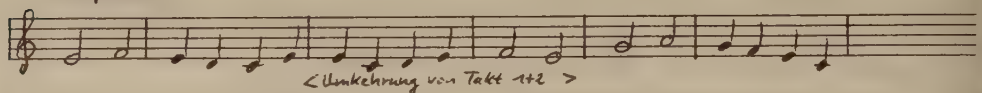


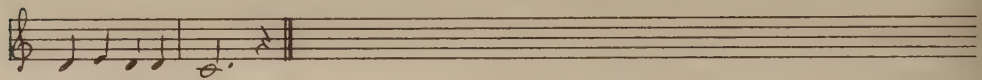
Beispiel 16: Heute ist im Schloß ein großer Ball (≙ B6)



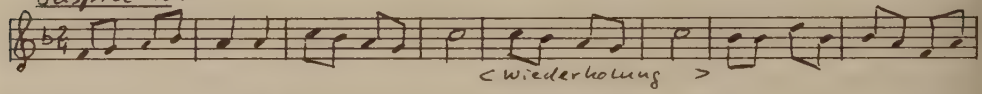


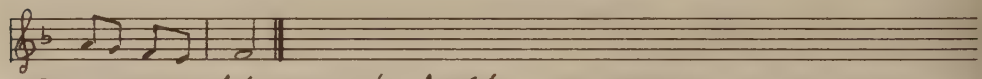
Beispiel 17:



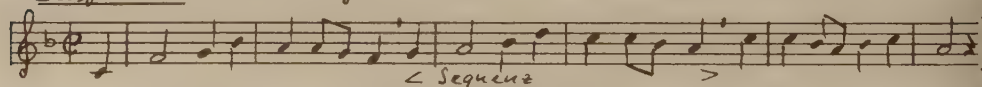


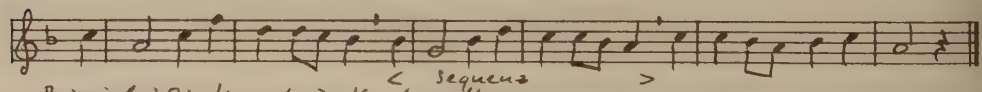
Beispiel 18:



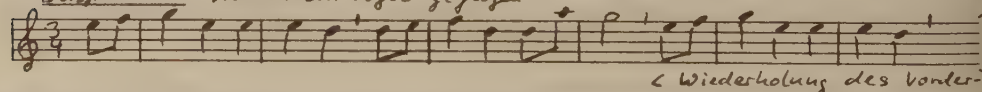


Beispiel 19: Ade zur guten Nacht

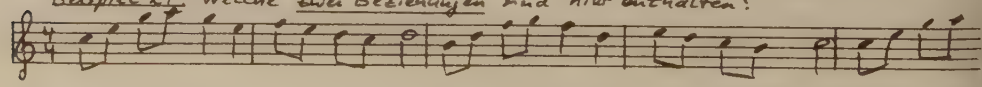


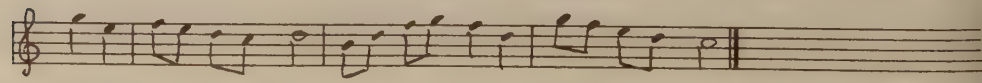


Beispiel 20: Kommt ein Vogel geflogen



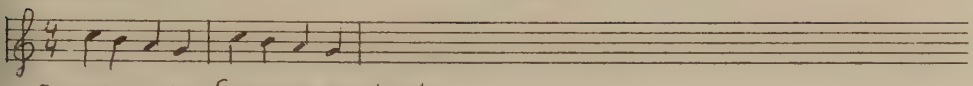
Beispiel 21: Welche zwei Beziehungen sind hier enthalten!





Antwort:

Beispiel 22: Wiederholung

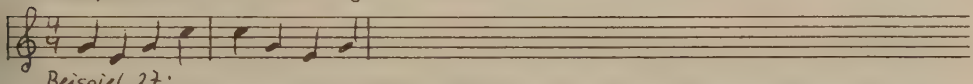


Beispiel 23: Sequenz nach oben

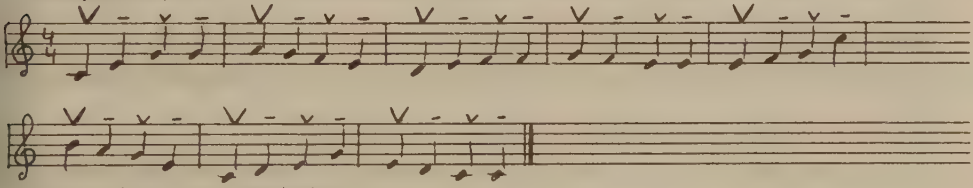
Beispiel 24: Sequenz nach unten

Beispiel 25: Umkehrung

Beispiel 26: Umkehrung



Beispiel 27:

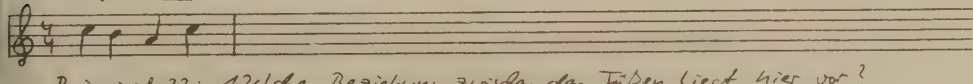


Beispiel 28: Wiederholung

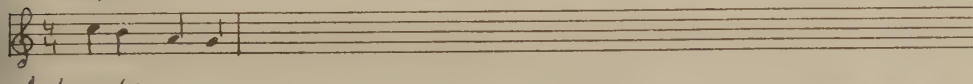
Beispiel 29: Sequenz

Beispiel 30: Umkehrung

Beispiel 31: keine Beziehung



Beispiel 32: Welche Beziehung zwischen den Tönen liegt hier vor?



Antwort:

Anhang

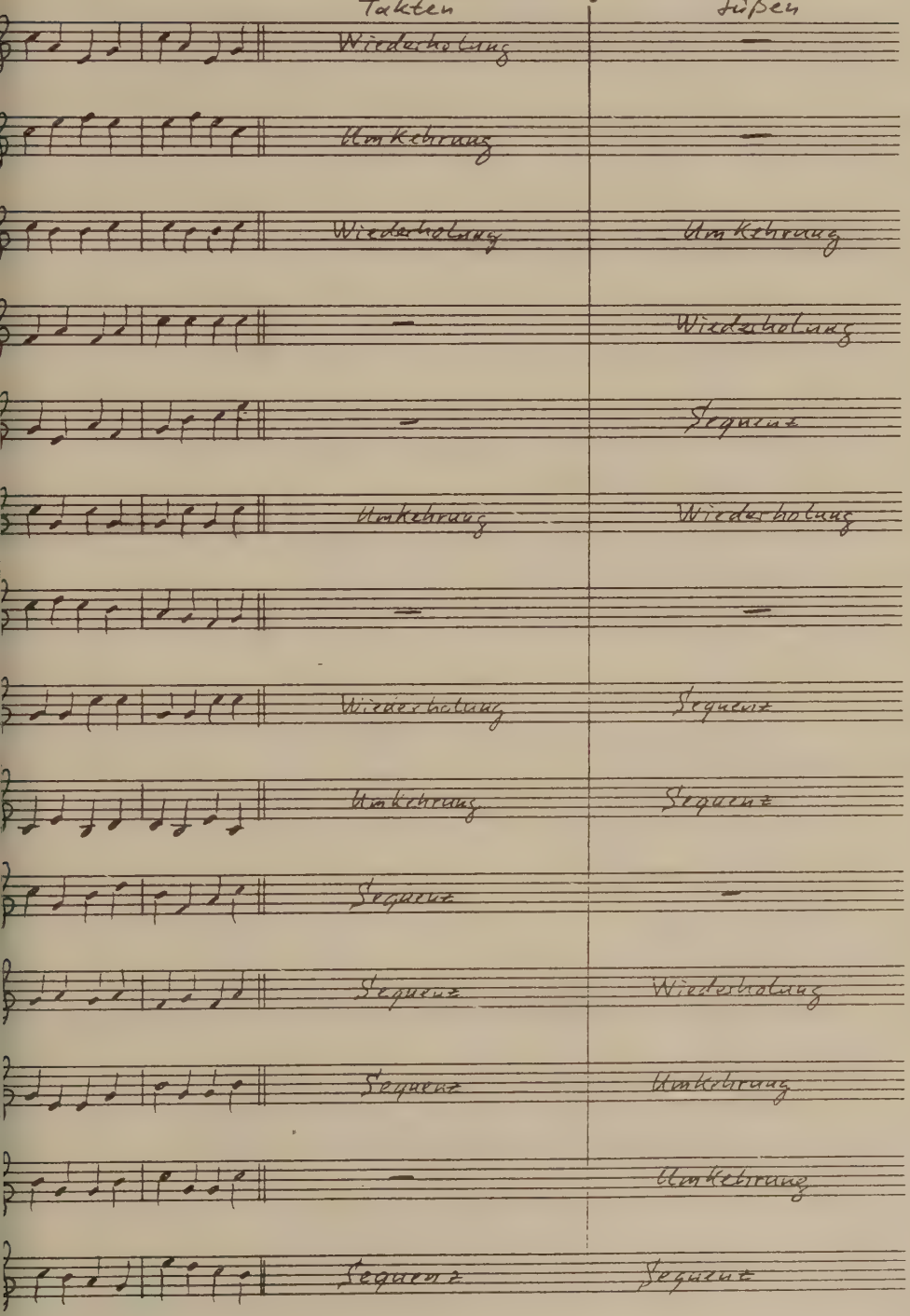
Beziehung zwischen

Nr.	Texten			Füßen		
	Wiederholung	Sequenz	Umkehrung	Wiederholung	Sequenz	Umkehrung
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						

Melodien für die 1. Hörübung in der 2. Sitzung

Beziehung zwischen

Takten	Füßen
Wiederholung	—
Umkehrung	—
Wiederholung	Umkehrung
—	Wiederholung
—	Sequenz
Umkehrung	Wiederholung
—	—
Wiederholung	Sequenz
Umkehrung	Sequenz
Sequenz	—
Sequenz	Wiederholung
Sequenz	Umkehrung
—	Umkehrung
Sequenz	Sequenz



Test zur Prüfung am Ende der zweiten Trainingssitzung

1. Handelt es sich bei der Melodie um den Fall einer Wiederholung des Vordersatzes mit Abwandlungen am Schluß?

   Ja / Nein

2. Welche Beziehung besteht bei der folgenden Melodie zwischen dem Vordersatzmotiv und dem Vordersatzanschlußglied?

   a) Sequenz
   b) Umkehrung
   c) keine
   d) Wiederholung

3. Ist bei der folgenden Melodie die Periode erweitert? Wenn ja, um welche Erweiterung handelt es sich?

   a) keine Erweiterung
   b) Erweiterung durch Motivsequenz
   c) Erweiterung durch Anschlußgliedwiederholung

4. Welche Beziehung besteht zwischen den beiden Takten des Vordersatzanschlußglieds?

   a) keine
   b) Wiederholung
   c) Sequenz
   d) Umkehrung

5. Welche Beziehung besteht zwischen den beiden Füßen des 1. Takt des Vordersatzmotivs?

   a) Wiederholung
   b) keine
   c) Umkehrung
   d) Sequenz

6. Liegt hier eine Erweiterung vor und wenn ja, um welche handelt es sich ?

   a) Erweiterung durch Anschlußgliedwiederholung
   b) keine Erweiterung
   c) Erweiterung durch Motivsequenz

7. Welche Beziehung besteht zwischen Nachsatzmotiv und Nachsatzanschlußglied?

   a) Umkehrung
   b) Wiederholung
   c) keine
   d) Sequenz

8. Wird bei dieser Melodie der Vordersatz mit Abwandlungen am Schluß als Nachsatz wiederholt?

   Ja / Nein

9. Welche Beziehung besteht hier zwischen den Füßen des zweiten Taktes des Nachsatzmotivs?

   a) keine
   b) Umkehrung
   c) Sequenz
   d) Wiederholung

10. Liegt bei der folgenden Melodie eine Erweiterung vor oder nicht und wenn ja, um welche Form der Erweiterung handelt es sich?

    a) Erweiterung durch Motivsequenz
    b) keine Erweiterung
    c) Erweiterung durch Anschlußgliedwiederholung

11. Welche Beziehung besteht zwischen den Takten des Nachsatzmotivs?

    a) Umkehrung
    b) keine
    c) Sequenz
    d) Wiederholung

12. Welche Beziehung besteht zwischen den Füßen des ersten Taktes des Vordersatzanschlußglieds?

    a) Wiederholung
    b) Sequenz
    c) keine
    d) Umkehrung

13. Welche Beziehung besteht zwischen den Takten des Nachsatzanschlußglieds?

a) Sequenz
b) Wiederholung
c) Umkehrung
d) keine

14. Welche Beziehung besteht bei der folgenden Melodie zwischen den Takten des Vordersatzanschlußglieds?

a) Umkehrung
b) keine
c) Sequenz
d) Wiederholung

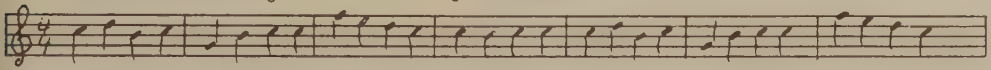
15. Welche Beziehung liegt bei der folgenden Melodie zwischen den Füßen des ersten Taktes des Vordersatzanschlußglieds vor?

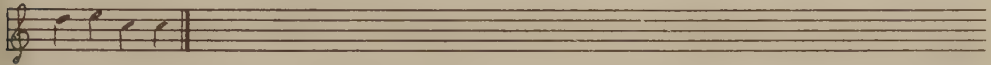
a) keine
b) Sequenz
c) Wiederholung
d) Umkehrung

Melodien für die zweite Hörübung am Ende der 2. Sitzung

♩ = 120

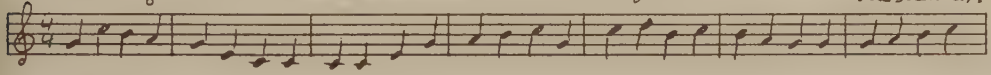
Vordersatzwiederholung mit Abwandlung am Schluß

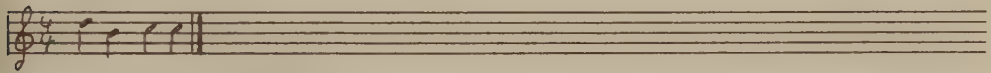




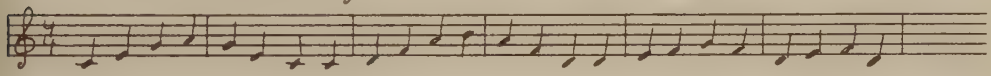
Umkehrung zwischen Vordersatzmotiv und -ausschlußglied

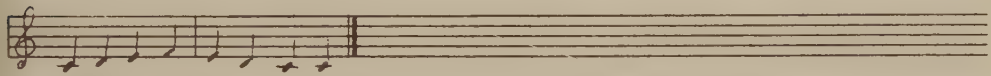
ausblenden!



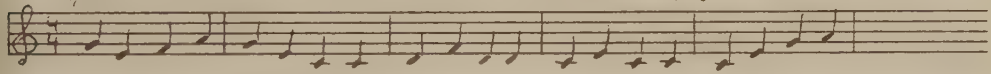


Periode ohne Erweiterung

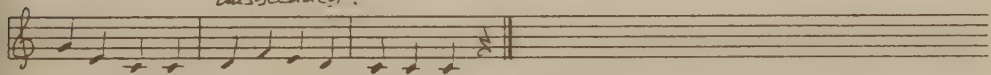




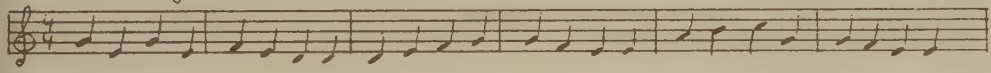
Sequenz zwischen den Takten des Vordersatzausschlußglieds

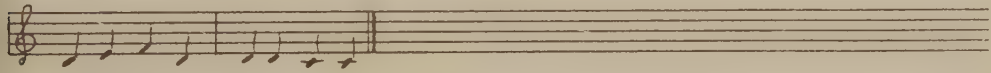


ausblenden!



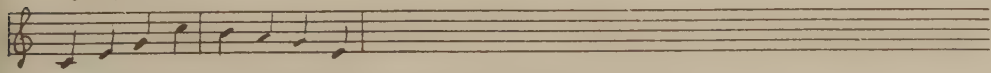
Wiederholung zwischen den Füßen des 1. Taktes des Vordersatzmotivs



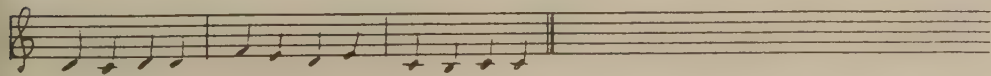


Periodenerweiterung durch Motivsequenz

ausblenden!

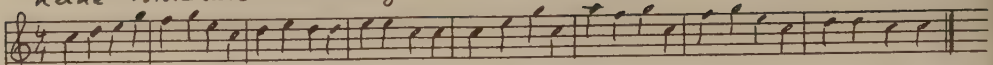


Sequenz zwischen Nachsatzmotiv und -ausschlußglied



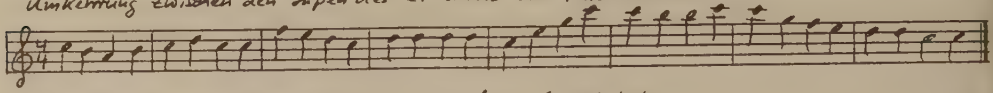
keine Vordersatzwiederholung

8.



Umkehrung zwischen den Füßen des 2. Taktes des Nachsatzmotivs

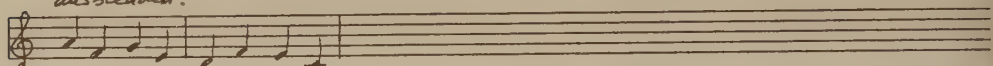
9.



Vordersatzerweiterung durch Anschlußgliedwiederholung

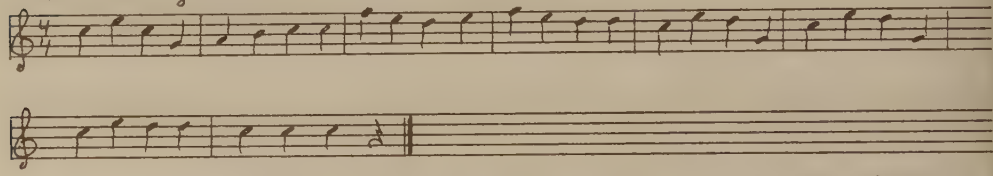
10.

ausblenden!



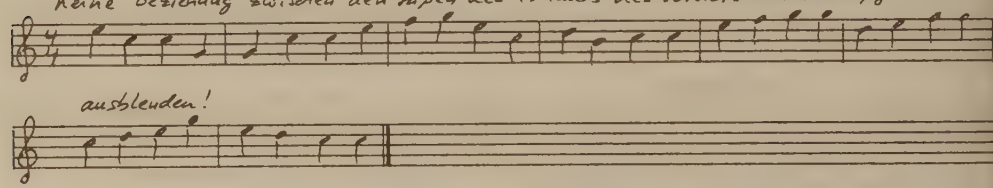
Wiederholung zwischen den Takten des Nachsatzmotivs

11.



keine Bezeichnung zwischen den Füßen des 1. Takts des Vordersatzanschlußglieds

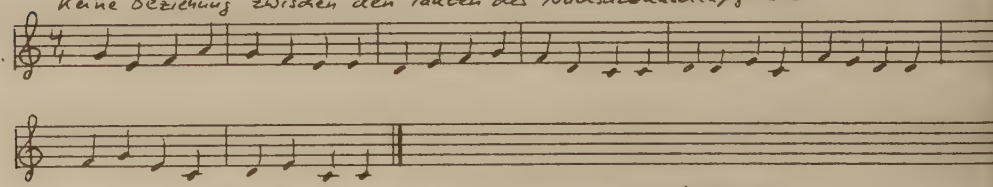
12.



ausblenden!

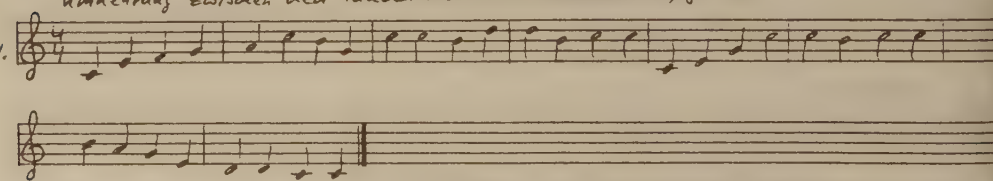
keine Bezeichnung zwischen den Takten des Nachsatzanschlußglieds

13.



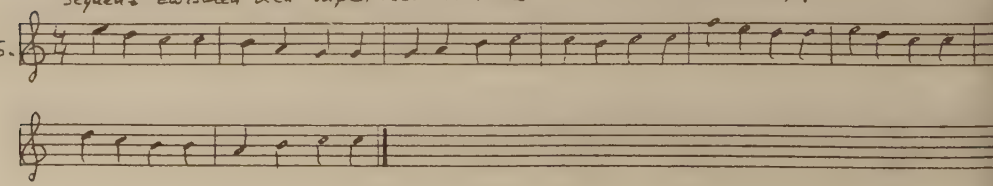
Umkehrung zwischen den Takten des Vordersatzanschlußglieds

14.

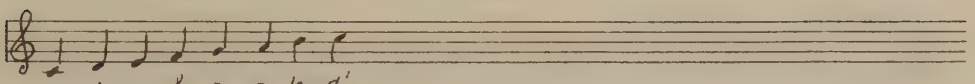


Sequenz zwischen den Füßen des 1. Takts des Vordersatzanschlußglieds

15.

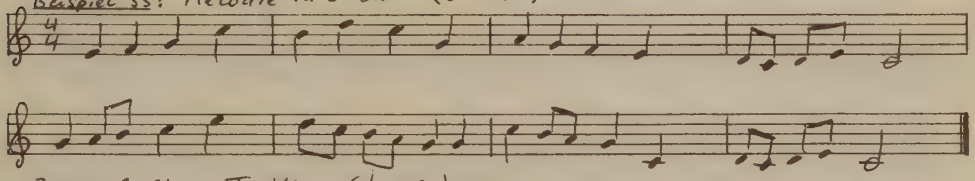


C-Dur Tonleiter

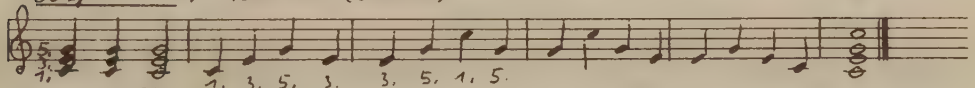


c d e f g a h c'

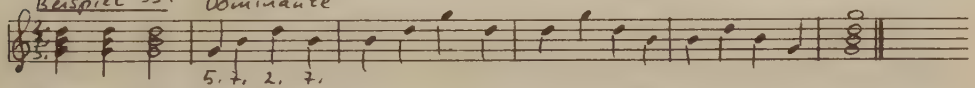
Beispiel 33: Melodie in C-Dur (♩ = 90)



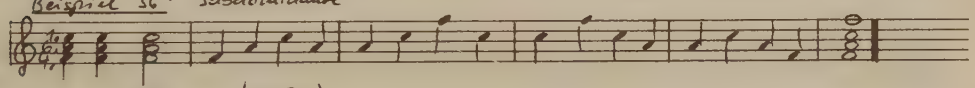
Beispiel 34: Tonika (♩ = 120)



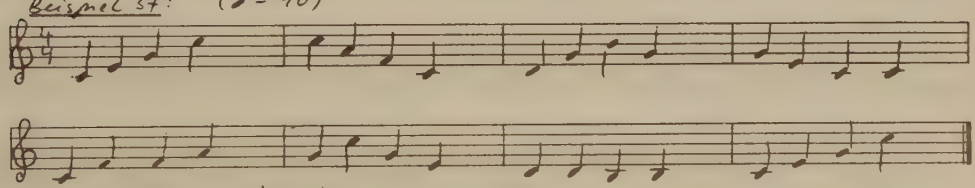
Beispiel 35: Dominante



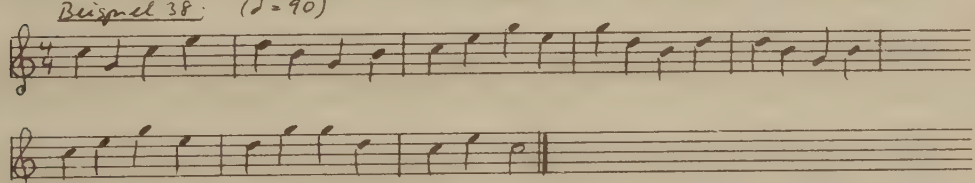
Beispiel 36: Subdominante



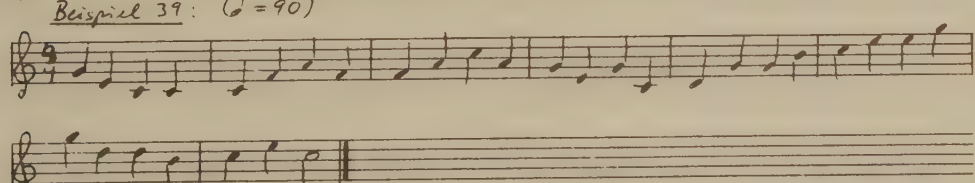
Beispiel 37: (♩ = 90)



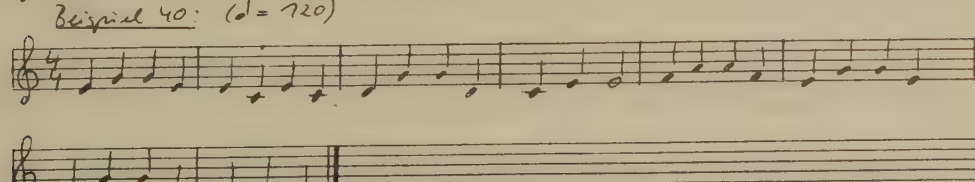
Beispiel 38: (♩ = 90)



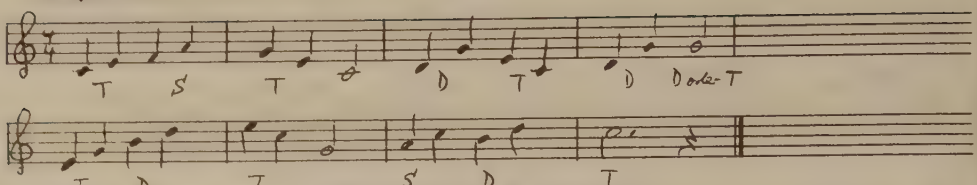
Beispiel 39: (♩ = 90)



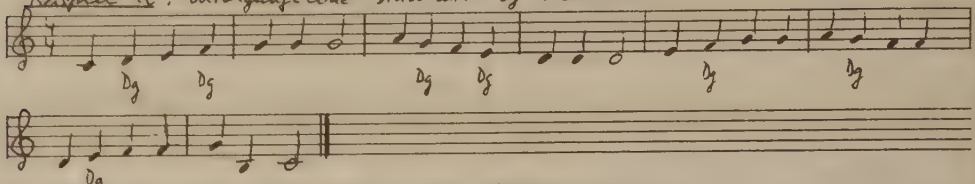
Beispiel 40: (♩ = 120)



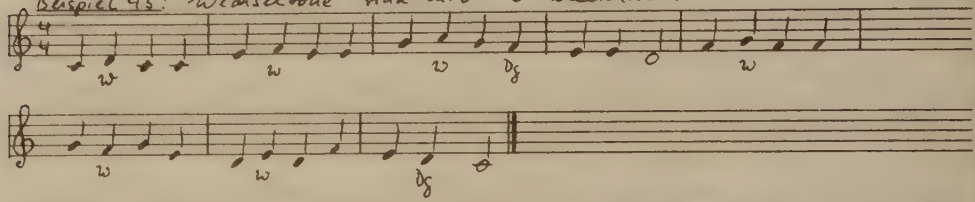
Beispiel 41 (♩ = 120)

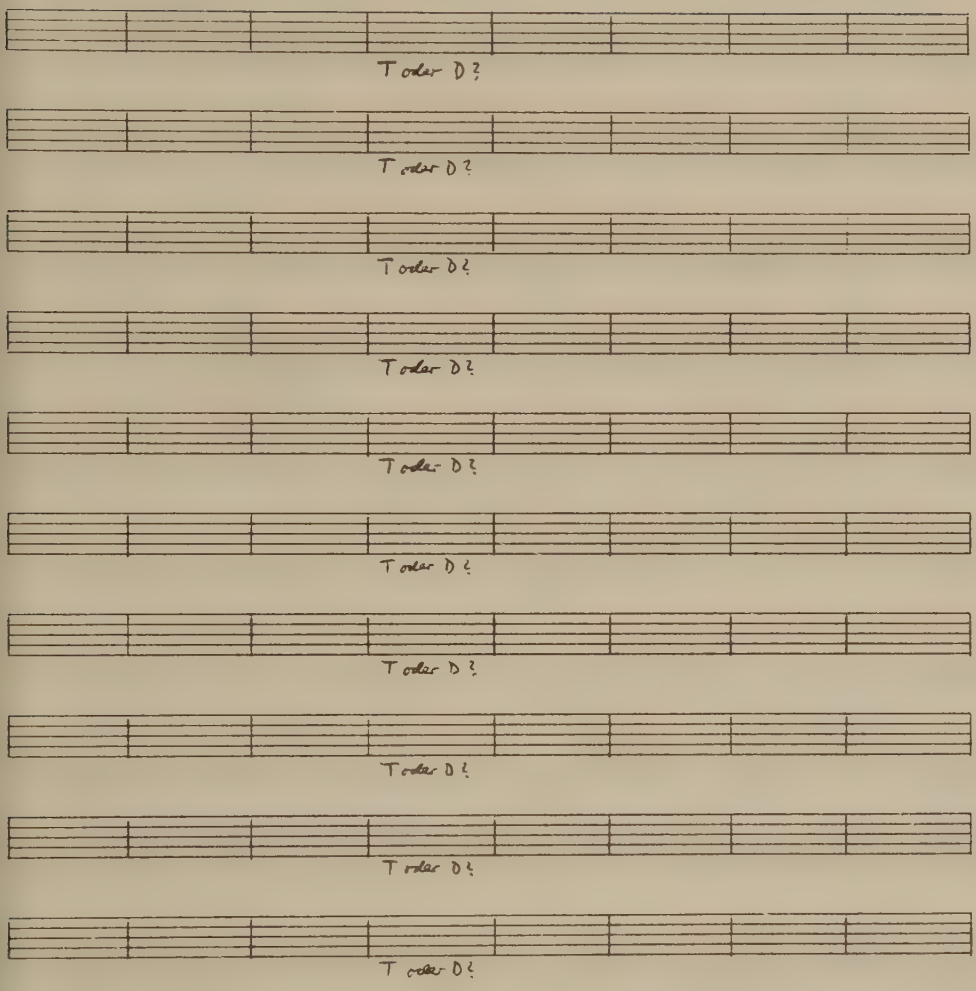


Beispiel 42: Durchgangstöne sind mit "Dg" bezeichnet!

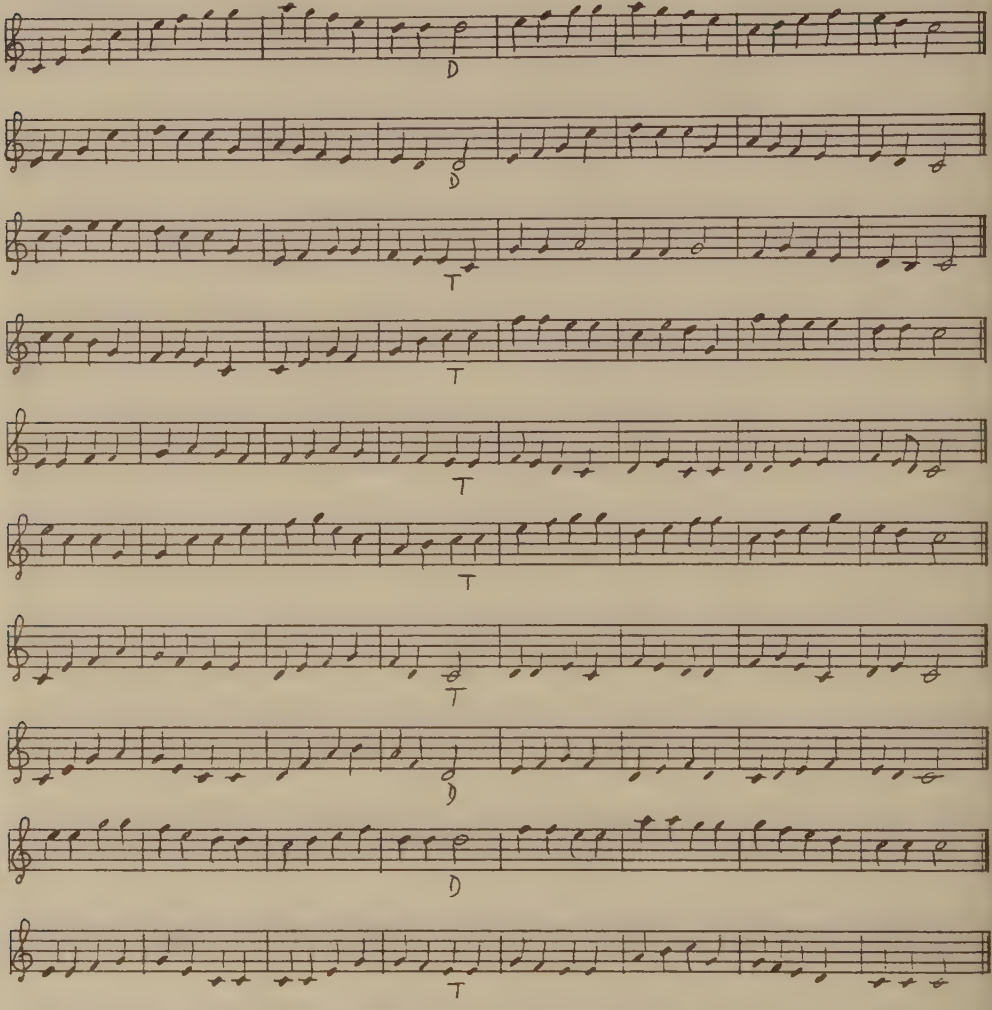


Beispiel 43: Wechseltöne sind mit "W" bezeichnet!




T oder D ?
T oder D ?
T oder D ?
T oder D ?
T oder D ?
T oder D ?
T oder D ?
T oder D ?
T oder D ?
T oder D ?

Melodien für die Hörübung "Konturzeichnen" am Anfang der 3. Sitzung
(jeweils 4/4 Takt, ♩ = 90)



Fragen zur Hörübung am Ende der dritten Trainingssitzung

1. Welche Beziehung besteht bei der folgenden Melodie zwischen Vordersatzmotiv und Vordersatzanschlußglied?

   a) Umkehrung
   b) keine
   c) Wiederholung
   d) Sequenz

2. Beginnt der Vordersatz auf dem Grundton oder einem anderen Ton der Tonika?

   a) auf einem anderen Ton der Tonika
   b) auf dem Grundton

3. Liegt bei der folgenden Melodie eine Erweiterung der Periode vor und wenn ja, um welche handelt es sich?

   a) Erweiterung durch Motivsequenz
   b) keine Erweiterung
   c) Erweiterung durch Anschlußgliedwiederholung

4. Endet der Vordersatz auf der Tonika oder auf der Dominate?

   a) auf der Dominante
   b) auf der Tonika

5. Welche Beziehung besteht zwischen den beiden Takten des Vordersatzanschlußglieds?

   a) Sequenz
   b) Wiederholung
   c) Umkehrung
   d) keine

6. Wird bei der folgenden Melodie der Vordersatz als Nachsatz mit Abwandlung am Schluß wiederholt?

   a) ja
   b) nein

7. Welche Beziehung besteht zwischen den beiden Füßen des zweiten Takts des Vordersatzmotivs?

   a) Wiederholung
   b) Umkehrung
   c) keine
   d) Sequenz

9. Endet der Vordersatz auf der Tonika oder auf der Dominante?

a) auf der Dominante

b) auf der Tonika

10. Endet der Nachsatz auf dem Grundton oder auf einem anderen Ton der Tonika?

a) auf dem Grundton
b) auf einem anderen Ton der Tonika

10. Welche Beziehung besteht zwischen den beiden Takten des Vordersatzmotivs?

a) Umkehrung
b) keine
c) Wiederholung
d) Sequenz

11. Welche Beziehung besteht zwischen den beiden Füßen des ersten Takts des Vordersatzanschlußglieds?

a) keine
b) Sequenz
c) Wiederholung
d) Umkehrung

12. Welche Beziehung besteht zwischen den beiden Takten des Nachsatzmotivs?

a) Sequenz
b) Wiederholung
c) Umkehrung
d) keine

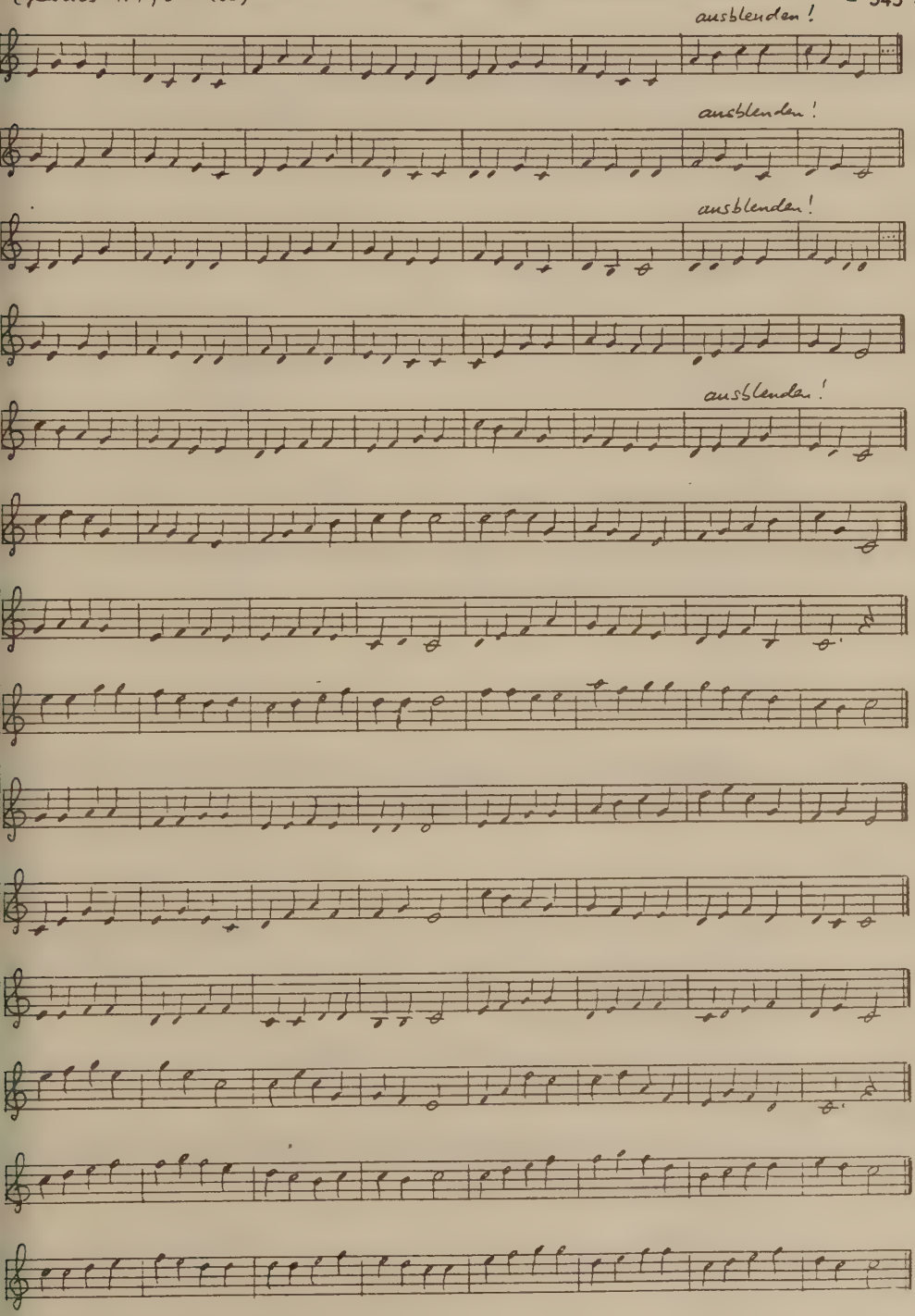
13. Welche Beziehung besteht zwischen Vorder- und Nachsatz?

a) keine
b) Wiederholung mit Abwandlung am Schluß

14. Welche Beziehung besteht zwischen Vordersatzmotiv und Vordersatzanschlußglied?

a) Sequenz
b) Umkehrung
c) keine
d) Wiederholung

Melodien für die Mörübung am Ende der dritten Trainingssitzung
(jeweils 4/4, ♩ = 120)



Fragen zur Hörübung in der vierten Trainingseinheit

1. Welche Beziehung besteht zwischen Vordersatzmotiv und Vordersatzanschlußglied?
a) keine
b) Sequenz
c) Wiederholung
d) Umkehrung

2. Welche Beziehung besteht zwischen den beiden Takten des Vordersatzmotivs?
a) Sequenz
b) Umkehrung
c) keine
d) Wiederholung

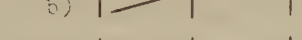
3. Welche Beziehung besteht zwischen den beiden Füßen im 1. Takt des Vordersatzmotivs?
a) Sequenz
b) keine
c) Umkehrung
d) Wiederholung

4. Wird bei der folgenden Melodie der Vordersatz als Nachsatz mit Abwandlung am Schluß wiederholt?
a) ja
b) nein

5. Liegt bei der folgenden Melodie eine erweiterte Periode vor und wenn ja, um welche Form der Erweiterung handelt es sich ?
a) keine Erweiterung
b) Erweiterung durch Anschlußgliedwiederholung
c) Erweiterung von Vorder- und Nachsatz durch Motivsequenz

6. Welche Beziehung besteht zwischen den Füßen im 2. Takt des Nachsatzmotivs?
a) keine
b) Umkehrung
c) Wiederholung
d) Sequenz

7. Welcher der vier Melodieverläufe paßt auf den Vordersatz (für zwei Takte jeweils ein Strich)?

a)
b)
c) 
d)

8. Welche Beziehung besteht zwischen Nachsatzmotiv und Nachsatzanschlußglied?
a) Umkehrung
b) keine
c) Sequenz
d) Wiederholung

9. Welche Beziehung besteht zwischen den Takten im Nachsatzmotiv?
a) keine
b) Umkehrung
c) Sequenz
d) Wiederholung

10. Welche Beziehung besteht zwischen Vordersatzmotiv und Vordersatzanschlußglied?
a) Sequenz
b) Wiederholung
c) Umkehrung
d) keine

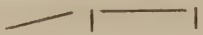
11. Endet der Vordersatz auf der Tonika oder der Dominante?
a) Dominante
b) Tonika

12. Welche Beziehung besteht zwischen den Füßen des 2. Takts des Vordersatzanschlußglieds?
a) Umkehrung
b) Wiederholung
c) Sequenz
d) keine

13. Wird bei der folgenden Melodie der Vordersatz als Nachsatz mit Abwandlung am Schluß wiederholt?

a) ja
b) nein

14. Welcher der vier Melodieverläufe paßt auf die beiden Takte des Nachsatzmotivs (pro Takt ein Strich)?

a)
b)
c) 
d)

15. Welche Beziehung besteht zwischen dem Nachsatzmotiv und dem Nachsatzanschlußglied?

a) Umkehrung
b) Sequenz
c) keine
d) Wiederholung

16. Welche Beziehung besteht zwischen den Füßen des 2. Takts des Vordersatzmotivs?

a) Sequenz
b) Wiederholung
c) keine
d) Umkehrung

17. Welche Beziehung besteht zwischen den Takten des Vordersatzanschlußglieds?

a) keine
b) Umkehrung
c) Sequenz
d) Wiederholung

18. Welcher der vier Melodieverläufe paßt auf den Nachsatz (für zwei Takte jeweils ein Strich)?

a)
b)
c)
d)

19. Liegt bei der folgenden Melodie eine erweiterte Periode vor und wenn ja, um welche Art der Erweiterung handelt es sich ?

a) Erweiterung durch Anschlußgliedwiederholung
b) keine Erweiterung
c) Erweiterung von Vorder- und Nachsatz durch Motivsequenz

20. Welche Beziehung besteht zwischen den beiden Füßen des 1. Takts des Nachsatzmotivs?

a) keine
b) Sequenz
c) Umkehrung
d) Wiederholung

21. Welche Beziehung besteht zwischen den beiden Takten des Nachsatzmotivs?

a) Wiederholung
b) Umkehrung
c) Sequenz
d) keine

22. Welche Beziehung besteht zwischen den beiden Füßen des 1. Takts des Vordersatzanschlußglieds?

a) Wiederholung
b) keine
c) Umkehrung
d) Sequenz

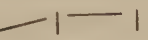
23. Welche Beziehung besteht zwischen dem Vordersatzmotiv und dem Vordersatzanschlußglied?

a) Umkehrung
b) keine
c) Wiederholung
d) Sequenz

24. Endet der Vordersatz auf der Tonika oder der Dominante?

a) Tonika
b) Dominante

25. Welcher der vier Melodieverläufe paßt auf den Nachsatz (für zwei Takte jeweils ein Strich)?

a) 

b)

c)

d)

26. Welche Beziehung besteht zwischen den beiden Füßen des 1. Takts des Nachsatzanschlußglieds?

a) Sequenz
b) Umkehrung
c) keine
d) Wiederholung

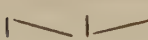
27. Welche Beziehung besteht zwischen den beiden Takten des Vordersatzmotivs?

a) Sequenz
b) keine
c) Umkehrung
d) Wiederholung

28. Liegt bei der folgenden Melodie eine erweiterte Periode vor und wenn ja, um welche Form der Erweiterung handelt es sich?

a) Erweiterung von Vorder- und Nachsatz durch Motivsequenz
b) Erweiterung des Vordersatzes durch Anschlußgliedwiederholun
c) keine Erweiterung

29. Welcher der vier Melodieverläufe paßt auf das Vordersatzanschlußglied (pro Takt ein Strich)?

a) 

b)

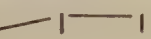
c)

d)

30. Welche Beziehung besteht zwischen den beiden Füßen des 2. Takts des Nachsatzmotivs?
a) Wiederholung
b) Sequenz
c) keine
d) Umkehrung

31. Welche Beziehung besteht zwischen den beiden Takten des Vordersatzmotivs?
a) Umkehrung
b) keine
c) Sequenz
d) Wiederholung

32. Welche Beziehung besteht zwischen dem Vordersatzmotiv und dem Vordersatzanschlußglied?
a) Wiederholung
b) Umkehrung
c) keine
d) Sequenz

33. Welcher der vier Melodieverläufe paßt auf das Vordersatzmotiv (pro Takt ein Strich)?
a) 
b)
c)
d)

34. Welche Beziehung besteht zwischen den beiden Füßen des 1. Takts des Vordersatzanschlußglieds?
a) Wiederholung
b) keine
c) Umkehrung
d) Sequenz

35. Wird bei der folgenden Melodie der Vordersatz als Nachsatz mit Abwandlung am Schluß wiederholt?
a) nein
b) ja

36. Welche Beziehung besteht zwischen den beiden Takten des Nachsatzmotivs?
a) Wiederholung
b) Umkehrung
c) Sequenz
d) keine

37. Liegt bei der folgenden Melodie eine erweiterte Periode vor und wenn ja, um welche Form der Erweiterung handelt es sich ?
a) Erweiterung des Vordersatzes durch Anschlußgliedwiederholun
b) Erweiterung von Vorder- und Nachsatz durch Motivsequenz
c) keine Erweiterung

38. Welche Beziehung besteht zwischen den beiden Füßen des 2. Takts des Vordersatzmotivs?
a) Wiederholung
b) Sequenz
c) keine
d) Umkehrung

39. Endet der Vordersatz auf der Dominante oder der Tonika?
a) Dominante
b) Tonika

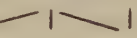
40. Welche Beziehung besteht zwischen dem Vordersatzmotiv und dem Vordersatzanschlußglied?
a) Sequenz
b) keine
c) Umkehrung
d) Wiederholung

41. Welche Beziehung besteht zwischen den beiden Takten des Vordersatzanschlußglieds?
a) Wiederholung
b) Umkehrung
c) keine
d) Sequenz

42. Welche Beziehung besteht zwischen den beiden Füßen des 2. Takts des Nachsatzmotivs?

a) Sequenz
b) Umkehrung
c) Wiederholung
d) keine

43. Welcher der vier Melodieverläufe paßt auf das Nachsatzanschlußglied (pro Takt ein Strich)?

a)
b)
c) 
d)

44. Liegt bei der folgenden Melodie eine erweiterte Periode vor und wenn ja, um welche Form der Erweiterung handelt es sich?

a) keine Erweiterung
b) Erweiterung durch Anschlußgliedwiederholung im Vordersatz
c) Erweiterung von Vorder- und Nachsatz durch Motivsequenz

45. Welche Beziehung Besteht zwischen den beiden Füßen des 1. Takts des Nachsatzmotivs?

a) keine
b) Wiederholung
c) Umkehrung
d) Sequenz

46. Welche Beziehung besteht zwischen den beiden Takten des Vordersatzmotivs?

a) Wiederholung
b) Umkehrung
c) Sequenz
d) keine

47. Welche Beziehung besteht zwischen Vorder- und Nachsatz?
a) keine
b) Wiederholung mit Abwandlung am Schluß

48. Welche Beziehung besteht zwischen den beiden Füßen des 2. Takts des Nachsatzanschlußglieds?
a) Sequenz
b) Umkehrung
c) keine
d) Wiederholung

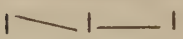
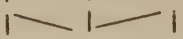
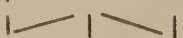
49. Endet der Vordersatz auf der Dominante oder der Tonika?
a) Tonika
b) Dominante

50. Welche Beziehung besteht zwischen den beiden Takten des Vordersatzanschlußglieds?
a) Umkehrung
b) keine
c) Wiederholung
d) Sequenz

51. Wird bei der folgenden Melodie der Vordersatz als Nachsatz mit Abwandlung am Schluß wiederholt?
a) nein
b) ja

52. Welche Beziehung besteht zwischen den beiden Füßen des 1. Takts des Vordersatzmotivs?
a) Umkehrung
b) Wiederholung
c) keine
d) Sequenz

53. Welcher der vier Melodieverläufe paßt auf das Nachsatzanschlußglied (pro Takt ein Strich)?

a) 
b) 
c) 
d)

54. Welche Beziehung besteht zwischen den beiden Füßen des 1. Takts des Nachsatzmotivs?
a) Umkehrung
b) Sequenz
c) Wiederholung
d) keine

55. Liegt bei der folgenden Melodie eine erweiterte Periode vor und wenn ja, um welche Form der Erweiterung handelt es sich?
a) Erweiterung von Vorder- und Nachsatz durch Motivsequenz
b) Erweiterung des Vordersatzes durch Anschlußgliedwiederholung
c) keine Erweiterung

56. Welche Beziehung besteht zwischen den beiden Füßen des 2. Takts des Vordersatzanschlußglieds?
a) keine
b) Umkehrung
c) Sequenz
d) Wiederholung

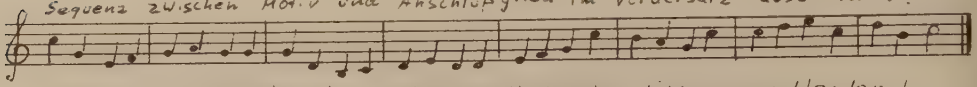
57. Welche Beziehung besteht zwischen dem Vordersatz und dem Nachsatz?
a) keine
b) Wiederholung mit Abwandlung am Schluß

58. Welche Beziehung besteht zwischen den beiden Füßen des Vordersatzanschlußglieds?
a) Wiederholung
b) keine
c) Sequenz
d) Umkehrung

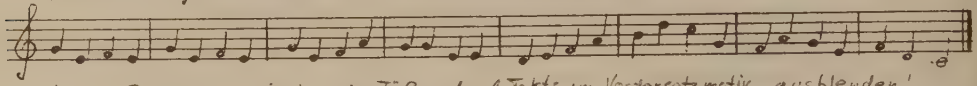
Abschlußtest für die vierte Trainingssitzung

(4/4 - Takt, ♩ = 120)

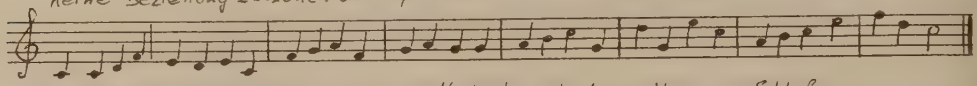
1. Sequenz zwischen Motiv und Anschlußglied im Vordersatz ausblenden!



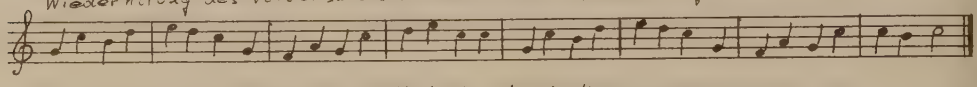
2. Wiederholung zwischen den Takten des Vordersatzmotivs ausblenden!



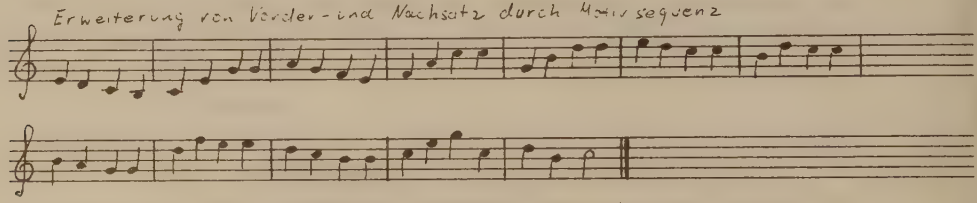
3. Keine Beziehung zwischen den Füßen des 1. Takts im Vordersatzmotiv ausblenden!



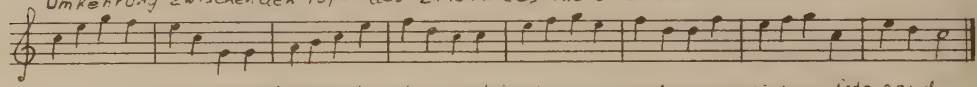
4. Wiederholung des Vordersatzes als Nachsatz mit Abwandlung am Schluß



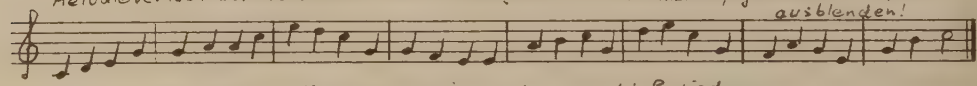
5. Erweiterung von Vorder- und Nachsatz durch Motivsequenz



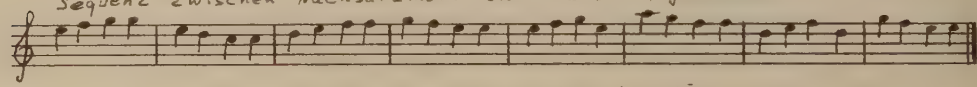
6. Umkehrung zwischen den Füßen des 2. Takts des Nachsatzmotivs



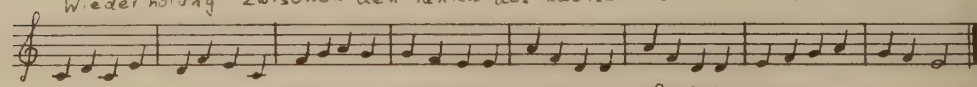
7. Melodieverlauf des Vordersatzmotivs absteigend und des Anschlußglieds aufsteigend ausblenden!



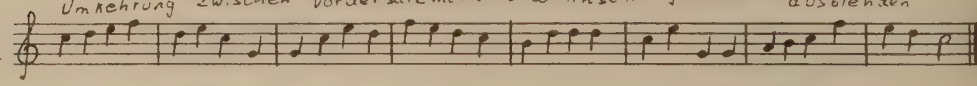
8. Sequenz zwischen Nachsatzmotiv und -anschlußglied



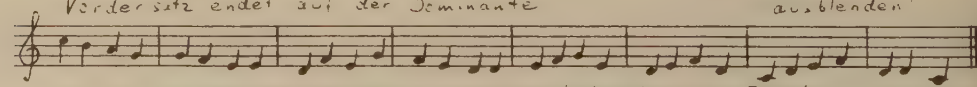
9. Wiederholung zwischen den Takten des Nachsatzmotivs



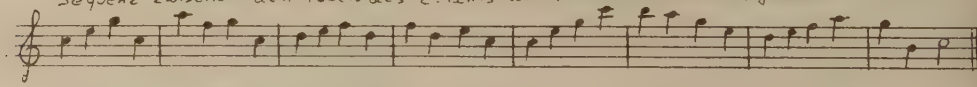
10. Umkehrung zwischen Vordersatzmotiv und -Anschlußglied ausblenden!



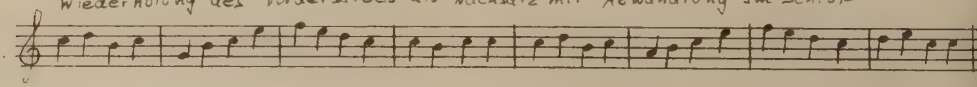
11. Vordersatz endet auf der Dominante ausblenden!



12. Sequenz zwischen den Füßen des 2. Takts der Vordersatzanschlußglieds

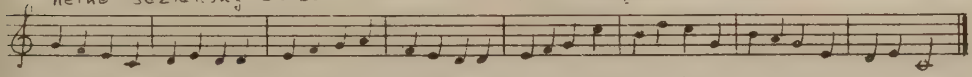


13. Wiederholung des Vordersatzes als Nachsatz mit Abwandlung am Schluß

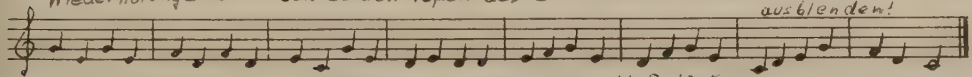


Melodieverlauf in den Takten des Nachsatzmotivs zuerst aufsteigend, dann absteigend

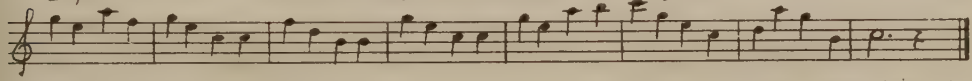
Keine Beziehung zwischen Motiv und Anschlußglied im Nachsatz



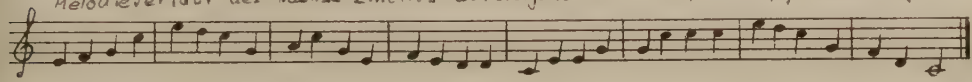
Wiederholung zwischen den beiden Füßen des 2. Takts des Vordersatzmotivs ausblenden!



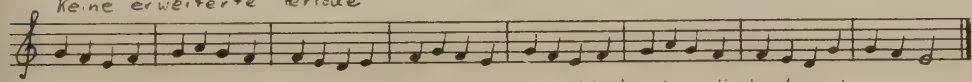
Sequenz zwischen den Takten des Vordersatzanschlußglieds ausblenden!



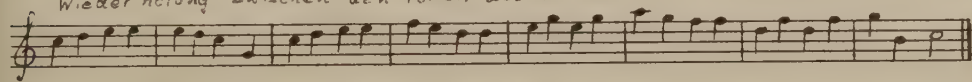
Melodieverlauf des Nachsatzmotivs aufsteigend und des Anschlußglieds absteigend



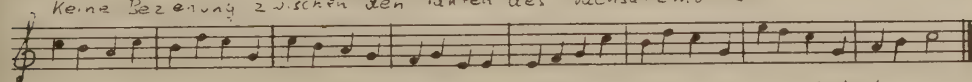
Keine erweiterte Periode



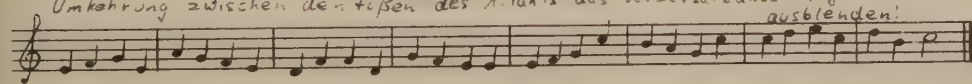
Wiederholung zwischen den Füßen des 1. Takts des Nachsatzmotivs



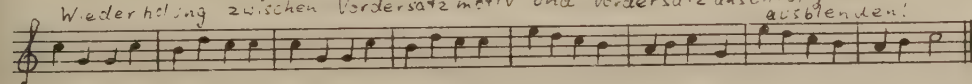
Keine Beziehung zwischen den Takten des Nachsatzmotivs



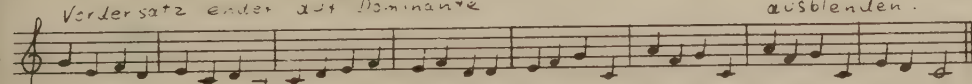
Umkehrung zwischen den Füßen des 1. Takts des Vordersatzanschlußglieds ausblenden!



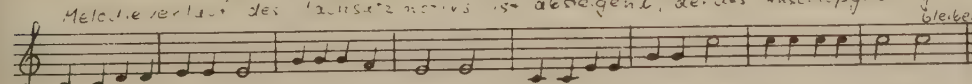
Wiederholung zwischen Vordersatzmotiv und Vordersatzanschlußglied ausblenden!



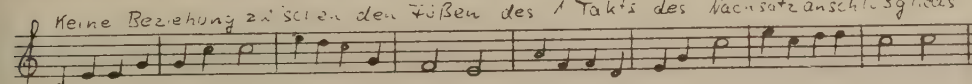
Vordersatz endet auf Dominante ausblenden!



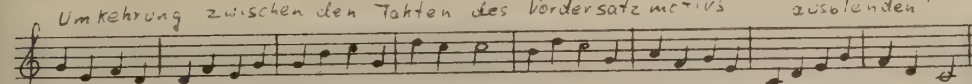
Melodieverlauf des Nachsatzmotivs absteigend, der des Anschlußglieds gleichbleibend



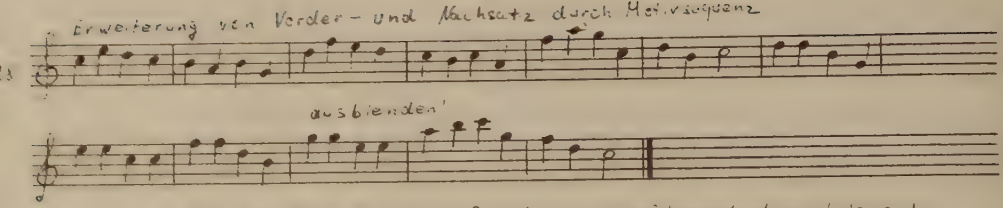
Keine Beziehung zwischen den Füßen des 1. Takts des Nachsatzanschlußglieds



Umkehrung zwischen den Takten des Vordersatzmotivs ausblenden!

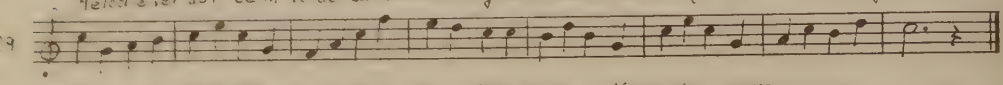


Erweiterung von Vorder- und Nachsatz durch Motivsequenz

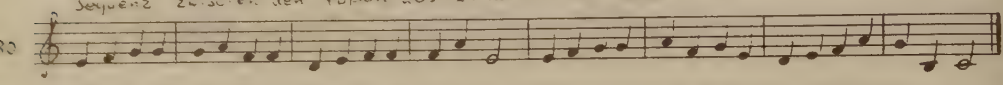


ausblenden!

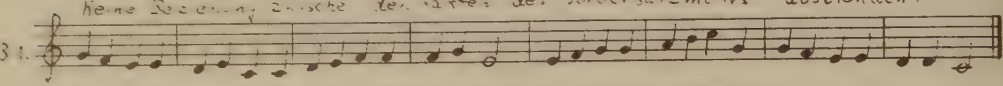
Melodieverlauf beim Vordersatzanschlußglied zuerst aufsteigend dann absteigend



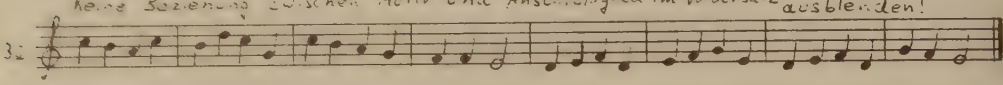
Sequenz zwischen den Füßen des 2. Takts des Nachsatzmotivs



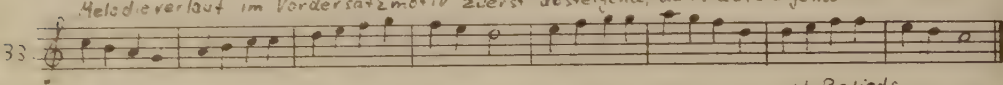
Keine Beziehung zwische den Takten des Vordersatzmotivs ausblenden!



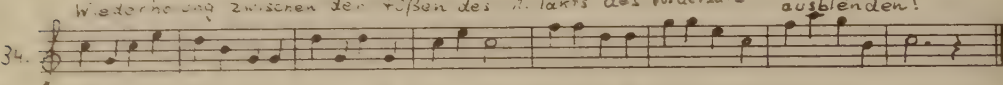
Keine Beziehung zwischen Motiv und Anschlußglied im Vordersatz ausblenden!



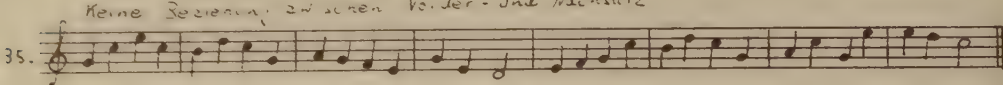
Melodieverlauf im Vordersatzmotiv zuerst absteigend, dann aufsteigend



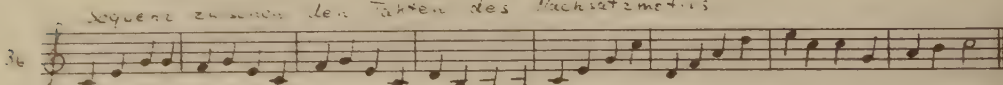
Wiederholung zwischen den Füßen des 1. Takts des Vordersatzanschlußglieds ausblenden!



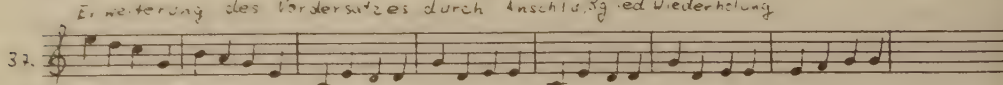
Keine Beziehung zwischen Vorder- und Nachsatz



Sequenz zwischen den Takten des Nachsatzmotivs

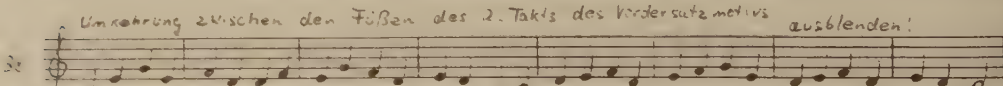


Erweiterung des Vordersatzes durch Anschlußglied Wiederholung

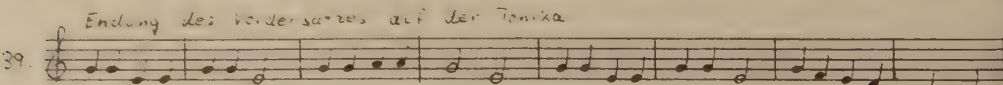


ausblenden!

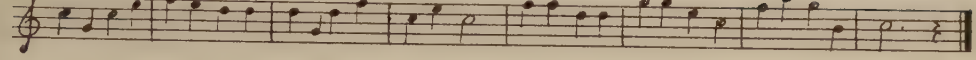
Umkehrung zwischen den Füßen des 2. Takts des Vordersatzmotivs ausblenden!



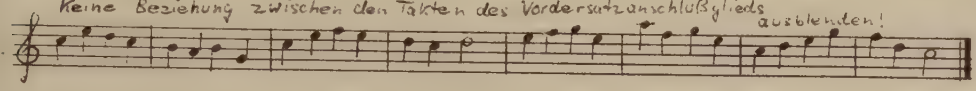
Endung des Vordersatzes auf der Tonika



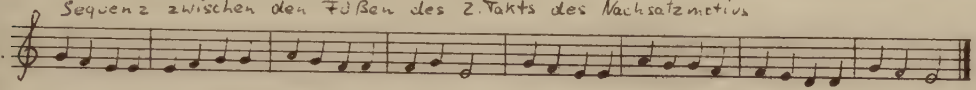
40 Keine Beziehung zwischen Vordersatz und -anschlußglied



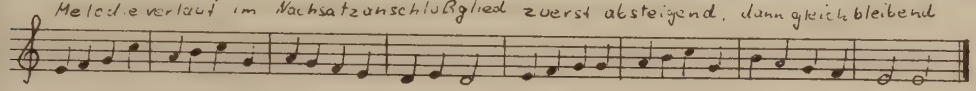
41. Keine Beziehung zwischen den Takten des Vordersatzanschlußglieds ausblenden!



42. Sequenz zwischen den Füßen des 2. Takts des Nachsatzmotivs

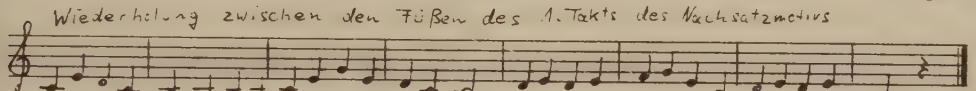


43 Melodieverlauf im Nachsatzanschlußglied zuerst absteigend, dann gleichbleibend

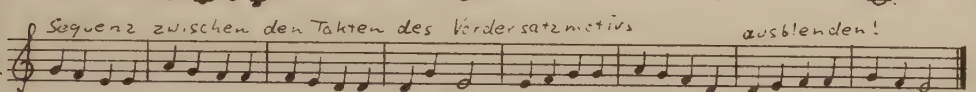


44. Keine Erweiterung der Periode

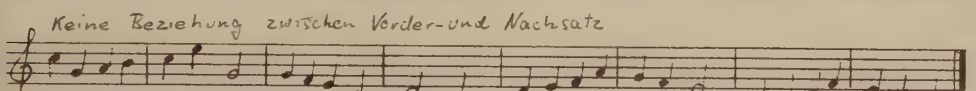
45. Wiederholung zwischen den Füßen des 1. Takts des Nachsatzmotivs



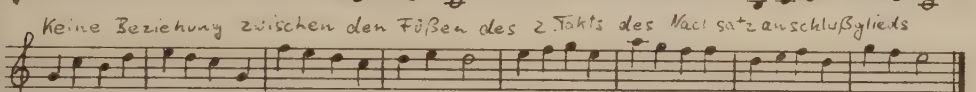
46. Sequenz zwischen den Takten des Vordersatzmotivs ausblenden!



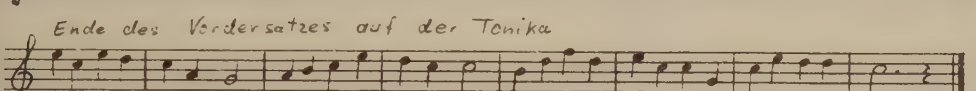
47. Keine Beziehung zwischen Vorder- und Nachsatz



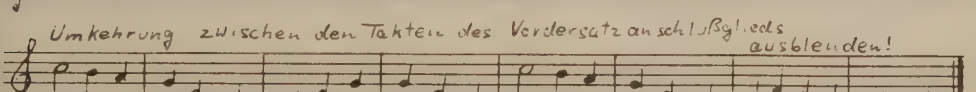
48. Keine Beziehung zwischen den Füßen des 2. Takts des Nachsatzanschlußglieds



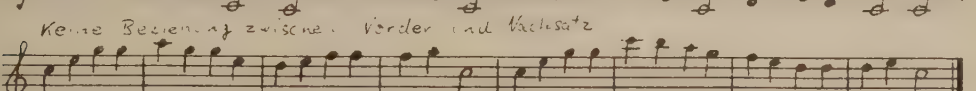
49. Ende des Vordersatzes auf der Tonika



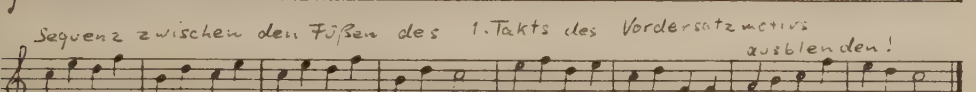
50. Umkehrung zwischen den Takten des Vordersatzanschlußglieds ausblenden!



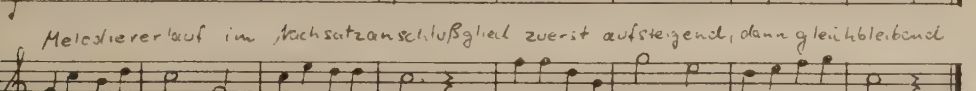
51. Keine Beziehung zwischen Vorder- und Nachsatz



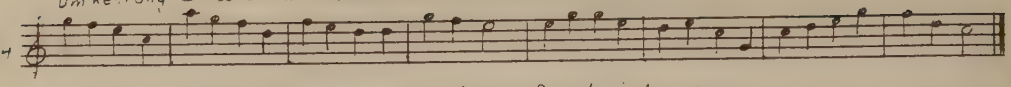
52. Sequenz zwischen den Füßen des 1. Takts des Vordersatzmotivs ausblenden!



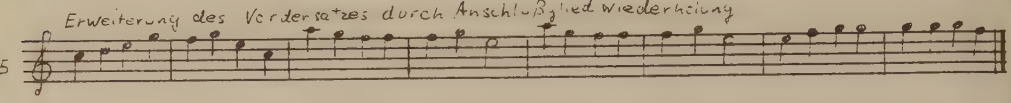
53. Melodieverlauf im Nachsatzanschlußglied zuerst aufsteigend, dann gleichbleibend



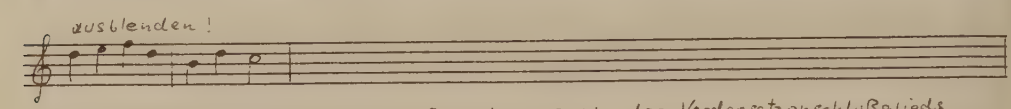
Umkehrung zwischen den Füßen des 1 Takts des Nachsatzmotivs



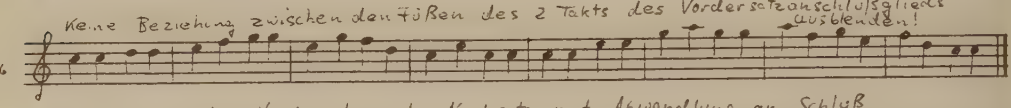
Erweiterung des Vordersatzes durch Anschlußglied Wiederholung



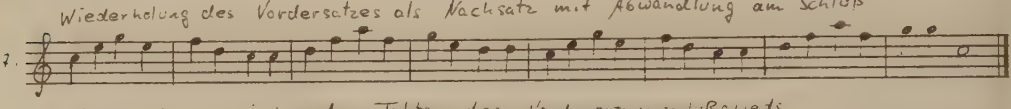
ausblenden!



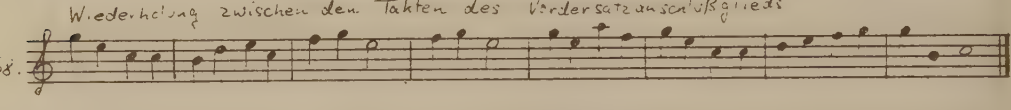
Keine Beziehung zwischen den Füßen des 2 Takts des Vordersatzanschlußglieds ausblenden!



Wiederholung des Vordersatzes als Nachsatz mit Abwandlung am Schluß



Wiederholung zwischen den Takten des Vordersatzanschlußglieds



ANHANG 3

Varianzanalysen

Anhang 3/1.1: Mit Kontrollbedingung als 5. Stufe des Faktors $SOA_{1,2}$

VARIANZ-URSACHE	S A Q	D F	M A Q	F-WERT	P %
ZWISCHEN VPN	1oo3266142.8	29			
1 f2,3	31926638.8	3	1o642212.9	3o.o8	o.o
FEHLER1	3o777654.o	87	353766.1		
2 f1,2	212o622o.8	6	353437o.1	34.26	o.o
FEHLER2	1795o914.3	174	1o3166.2		
3 SOA1,2	118798161.o	4	2969954o.3	5o.47	o.o
FEHLER3	68266395.5	116	5885o3.4		
4 SOA2,3	14239o88.o	1	14239o88.o	23.44	o.o
FEHLER4	17619o16.8	29	6o7552.3		
5 Konfiguration	29947911.8	1	29947911.8	47.86	o.o
FEHLER5	18146757.3	29	62575o.3		
1x2	18o7196.8	18	1oo399.8	3.88	o.o
FEHLER12	13518475.o	522	25897.5		
1x3	1481711.3	12	123475.9	5.16	o.o
FEHLER13	8323728.o	348	23918.8		
1x4	22o623.5	3	73541.2	3.33	2.32
FEHLER14	192o896.3	87	22o79.3		
1x5	1434688.3	3	478229.4	15.55	o.o
FEHLER15	2675o76.5	87	3o748.o		
2x3	6766521.3	24	281938.4	9.46	o.o
FEHLER23	2o745623.5	696	298o6.9		
2x4	66o66o.3	6	11o11o.o	4.61	o.o2
FEHLER24	4157139.5	174	23891.6		
2x5	97o8441.3	6	1618o73.5	27.33	o.o
FEHLER25	1o3o3449.8	174	59215.2		
3x4	157o639.8	4	392659.9	1o.51	o.o
FEHLER34	4333696.5	116	37359.5		
3x5	13859764.o	4	3464941.o	23.5o	o.o
FEHLER35	171o5429.3	116	14746o.6		
4x5	84818o.3	1	84818o.3	13.78	o.o9
FEHLER45	1785o59.8	29	61553.8		
1x2x3	2269126.o	72	31515.6	1.38	1.93
FEHLER123	47547594.o	2o88	22771.8		
1x2x4	9o4848.5	18	5o269.4	2.21	o.29
FEHLER124	11863923.3	522	22727.8		
1x2x5	1323761.o	18	73542.3	3.68	o.o
FEHLER125	1o56689o.3	522	2o243.1		
1x3x4	56o44o.o	12	467o3.3	2.12	1.54
FEHLER134	767o254.5	348	22o41.o		
1x3x5	499531.o	12	41627.6	2.35	o.67

Blatt 2 zu:
Anhang 3/1.1: Mit Kontrollbedingung als 5. Stufe des Faktors $SOA_{1,2}$

VARIANZ-URSACHE	S A Q	D F	M A Q	F-WERT	P %
FEHLER135	6176289.8	348	17748.o		
1x4x5	1oo425.5	3	33475.2	1.77	15.8o
FEHLER145	1641642.5	87	18869.5		
2x3x4	842549.7	24	351o6.2	1.51	5.66
FEHLER234	16189923.5	696	23261.4		
2x3x5	2557lo2.o	24	1o6545.9	3.99	o.o
FEHLER235	186o3644.o	696	26729.4		
2x4x5	158276.o	6	26379.3	1.18	32.18
FEHLER245	39o6259.5	174	22449.8		
3x4x5	1o932o8.o	4	2733o2.o	7.62	o.o
FEHLER345	4161284.3	116	35873.1		
1x2x3x4	189715o.5	72	26349.3	1.18	14.18
FEHLER1234	46487o31.3	2o88	22263.9		
1x2x3x5	2o18767.8	72	28o38.4	1.39	1.75
FEHLER1235	42o5o866.3	2o88	2o139.3		
1x2x4x5	5875oo.5	18	32638.9	1.48	9.32
FEHLER1245	11548251.o	522	22123.1		
1x3x4x5	265122.5	12	22o93.5	1.o1	43.71
FEHLER1345	7613359.8	348	21877.5		
2x3x4x5	6775o1.5	24	28229.2	1.4o	9.72
FEHLER2345	14o31395.o	696	2o16o.1		
1x2x3x4x5	2ooo251.5	72	27781.3	1.35	2.68
FEHLER12345	42818848.5	2o88	2o5o7.1		
INNERHALB VPN	8o2738777.8	1677o	47867.5		
GESAMT	18o6oo492o.5	16799	1o75o6.7		

Anhang 3/1.2: Ohne Kontrollbedingung

VARIANZ-URSACHE	S A Q	D F	M A Q	F-Wert	P %
ZWISCHEN VPN	1oo585o67o.o	29			
1 f2,3	3o387373.8	3	1o129124.6	29.49	o.o
FEHLER1	2988o951.o	87	343459.2		
2 f1,2	269o8981.8	6	448483o.3	35.98	o.o
FEHLER2	21719676.o	174	124825.7		
3 SOA1,2	2917435.8	3	972478.6	18.34	o.o
FEHLER3	4611997.o	87	53o11.5		
4 SOA2,3	15324145.5	1	15324145.5	24.91	o.o
FEHLER4	17839649.o	29	61516o.3		
5 Konfigu-ration	42369383.8	1	42369383.8	74.53	o.o
FEHLER5	16485598.3	29	568468.9		
1 x 2	1695616.3	18	942oo.9	3.52	o.o
FEHLER12	13967o69.3	522	26756.8		
1x3	239731.5	9	26636.8	1.34	21.74
FEHLER13	519686o.3	261	19911.3		
1x4	315924.o	3	1o53o8.o	3.92	1.12
FEHLER14	23359o8.o	87	26849.5		
1x5	17o7838.3	3	569279.4	2o.88	o.o
FEHLER15	2371445.5	87	27258.o		
2x3	6o59o3.8	18	33661.3	1.33	16,35
FEHLER23	13221o64.8	522	25327.7		
2x4	75o596.5	6	125o99.4	5.o7	o.o1
FEHLER24	4296olo.8	174	24689.7		
2x5	111248o7.3	6	1854134.5	26.53	o.o
FEHLER25	12161923.o	174	69896.1		
3x4	115278.5	3	38426.2	1.3o	25.92
FEHLER34	2569428.3	87	29533.7		
3x5	827483.5	3	275827.8	9.24	o.o
FEHLER35	2596523.8	87	29845.1		
4x5	1114322.3	1	1114322.3	15.37	o.o
FEHLER45	21o3o45.5	29	72518.8		
1x2x3	169o564.3	54	313o6.7	1.31	6.67
FEHLER123	37412785.3	1566	2389o.7		
1x2x4	697865.7	18	3877o.3	1,68	3.82
FEHLER124	12o14962.5	522	23o17.2		
1x2x5	1659578.5	18	92198.8	4,o5	o.o
FEHLER125	11886282.o	522	2277o.7		
1x3x4	449694.8	9	49966.1	2.57	o.74
FEHLER134	5o65349.o	261	194o7.5		
1x3x5	156677.5	9	174o8.6	1.o2	42.16

Blatt 2 zu:
Anhang 3/1.2: Ohne Kontrollbedingung

VARIANZ-URSACHE	S A Q	D F	M A Q	F-Wert	P %
FEHLER135	444o334.5	261	17o12.8		
1x4x5	61o28.5	3	2o342.8	o.87	46.o8
FEHLER145	2o38447.o	87	2343o.4		
2x3x4	448953.o	18	24941.8	1.o9	35.57
FEHLER234	11915265.o	522	22826.2		
2x3x5	899486.o	18	49971.4	2.oo	o.87
FEHLER235	13o74324.5	522	25o46.6		
2x4x5	115423.3	6	19237.2	o.83	55.1o
FEHLER245	4o51397.3	174	23283.9		
3x4x5	824374.o	3	274791.3	7.95	o.o1
FEHLER345	3oo8332.3	87	34578.5		
1x2x3x4	1675998.8	54	31o37.o	1.33	5.61
FEHLER1234	36511635.3	1566	23315.2		
1x2x3x5	13178o6.5	54	244o3.8	1.23	12.59
FEHLER1235	311o77o7.3	1566	19864.4		
1x2x4x5	599766.5	18	3332o.4	1,59	5.84
FEHLER1245	1o953847.3	522	2o984.4		
1x3x4x5	241263.3	9	268o7.o	1.29	24.21
FEHLER1345	54225o1.8	261	2o775.9		
2x3x4x5	587314.8	18	32628.6	1.58	6.13
FEHLER2345	1o8o7417.o	522	2o7o3.9		
1x2x3x4x5	16oo629.8	54	29641.3	1.38	3.6o
FEHLER12345	33582497.o	1566	21444.8		
INNERHALB VPN	534o81482.3	1341o	39827.1		
GESAMT	1539932152.3	13439	114586.8		

Anhang 3/2

VARIANZ-URSACHE	S A Q	D F	M A Q	F-Wert	P %
ZWISCHEN VPN	374493o56.3	26			
1 Konfigu-	17o7o891.6	2	8535445.8	7o.61	o.o
FEHLER 1 rqtion	6285626.4	52	12o877.4		
2 Aufgabe	66o1331.6	2	33oo665.8	6.84	o.23
FEHLER 2	25o81547.2	52	482337.4		
3 same/diffe-	13345113.5	1	13345113.5	36.35	o.oo
FEHLER 3 rent	9546282.2	26	367164.7		
4 SOA	286696.4	1	286696.4	4.o8	5.38
FEHLER 4	18275o4.8	26	7o288.6		
5 f	4o395o3.8	1	4o395o3.8	35.7o	o.o
FEHLER 5	2942337.8	26	113166.8		
1X2	23o44o6.2	4	5761o1.6	5,29	o.o6
FEHLER 12	113336o8.5	1o4	1o8977.o		
1X3	585643.9	2	292821.9	2.72	7.53
FEHLER 13	56o2162.2	52	1o7733.9		
1X4	1o32619.2	2	5163o9.6	7.2o	o.17
FEHLER 14	373o14o.6	52	71733.5		
1X5	58o5955.o	2	29o2977.5	3o.5o	o.o
FEHLER 15	4948822.o	52	95169.7		
2X3	521853.7	2	26o926.9	2,28	11.27
FEHLER 23	5959196.7	52	114599.9		
2X4	567499o.5	2	2837495.2	29.72	o.o
FEHLER 24	4964466.8	52	9547o.5		
2X5	8514988.7	2	4257494.4	55.7o	o.o
FEHLER 25	3974596.5	52	76434.5		
3X4	12o637.1	1	12o637.1	5.65	2.5o
FEHLER 34	555116.7	26	2135o.6		
3X5	754336.7	1	754336.7	8.18	o.82
FEHLER 35	2397658.9	26	92217.7		
4X5	62o8.1	1	62o8.1	o.12	72.93
FEHLER 45	131945o.o	26	5o748.1		
1X2X3	954772.5	4	238693.1	2.48	4.81
FEHLER 123	99947o6.3	1o4	961o2.9		
1X2X4	827428.1	4	2o6857.o	3.26	1.45
FEHLER 124	6598765.7	1o4	63449.7		
1X2X5	2677419.3	4	669354.8	11.83	o.o
FEHLER 125	5885o76.5	1o4	56587.3		
1X3X4	82474.o	2	41237.o	o.88	42.oo

Blatt 2 zu:
Anhang 3/2

VARIANZ-URSACHE	S A Q	D F	M A Q	F-Wert	P %
FEHLER 134	2431129.8	52	46752.5		
1X3X5	5o781.o	2	2539o.5	o.43	65.13
FEHLER 135	3o54694.6	52	58744.1		
1X4X5	165848.8	2	82924.4	1.o1	37.29
FEHLER 145	42893o1.7	52	82486.6		
2X3X4	611966.o	2	3o5983.o	6.81	o,23
FEHLER 234	233625o.8	52	44927.9		
2X3X5	1671278.7	2	835639.4	15.o8	o.o1
FEHLER 235	28814o1.3	52	55411.6		
2X4X5	2388oo.6	2	1194oo.3	4.33	1.81
FEHLER 245	1433126.7	52	2756o.1		
3X4X5	386419.o	1	386419.o	8.35	o.76
FEHLER 345	12o3745.7	26	46297.9		
1X2X3X4	511864.o	4	127966.o	2.76	3.16
FEHLER 1234	4825654.7	1o4	464oo.5		
1X2X3X5	76o486.o	4	19o121.5	2.84	2.98
FEHLER 1235	6972284.8	1o4	67o41.2		
1X2X4X5	42o15o.7	4	1o5o37.7	1.83	12.89
FEHLER 1245	5975875.o	1o4	5746o.3		
1X3X4X5	28445.8	2	14222.9	o.25	78.32
FEHLER 1345	3o13o84.5	52	57943.9		
2X3X4X5	578989.8	2	289494.9	3.53	3.65
FEHLER 2345	4267o36.o	52	82o58.4		
1X2X3X4X5	115853.7	4	28963.4	o.6o	66.65
FEHLER 12345	5o56413.7	1o4	48619.4		
INNERHALB VPN	235355215.8	1917	122772.7		
GESAMT	611928277.2	1943	343869.4		

Anh. 3/3.1: Varianzanalyse der Lokalisationshäufigkeiten

VARIANZ-URSACHE	S A Q	D F	M A Q	F-WERT	P %
ZWISCHEN VPN	1.3	43			
1	137o4.7	16	856.5	1o2.69	o.o
FEHLER 1	5738.6	688	8.3		
INNERHALB VPN	19443.3	7o4	27.6		
GESAMT	19444.6	747	26.o		

Anh. 3/3.2: Varianzanalyse der Klickversichiebungsscores

VARIANZ-URSACHE	S A Q	D F	M A Q	F-Wert	P %
ZWISCHEN VPN	89.5	43			
1	374.3	16	23.4	21.o4	o.o
FEHLER 1	765.o	688	1.1		
INNERHALB VPN	1139.4	7o4	1.6		
GESAMT	1228.8	747	1.6		

Anh. 3/4

VARIANZURSACHE	S A Q	DF	M A Q	F-Wert	P %
ZWISCHEN VPN	28o466o.69	86	32612.33		
FAKTOR A: EXPERIMENT.VS. KONTR.-GR.	147986.o1	1	147986.o1	4.73	3.23
VPN INNERHALB DER GRUPPEN (FEHLER 1)	2656674.68	85	31255.oo		
INNERHALB VPN	1627oo8.13	13o5	1246.75		
FAKTOR B: KLICKPOSITIONEN 1 bis 16	827175.94	15	55145.o6	89.92	o.o
INTERAKTION AXB	17899.5o	15	1193.3o	1.95	1.61
INTERAKTION BX VPN INNERHALB GRUPPEN (FEHLER 2)	781932.68	1275	613.28		
GESAMT	4431668.82	1391	3185.96		

Anh. 3/5

VARIANZ-URSACHE	S A Q	D F	M A Q	F-WERT	P %
ZWISCHEN VPN	361539262.1	37			
1	17884445.3	2	8942222.6	5.43	o.63
FEHLER 1	121848166.1	74	1646596.8		
2	28643989.3	2	14321994.7	58.57	o.o
FEHLER 2	18o94874.o	74	244525.3		
3	8246726.6	1	8246726.6	15.84	o.o3
FEHLER 3	19257411.4	37	52o47o.6		
1X2	1565961.6	4	39149o.4	1.57	18.5o
FEHLER 12	3688o939.1	148	249195.5		
1X3	264o27o.2	2	132o135.1	4.o7	1.98
FEHLER 13	24124978.4	74	325o13.2		
2X3	982918.2	2	491459.1	2.16	12.23
FEHLER 23	168219o2.2	74	227323.o		
1X2X3	2241486.7	4	56o371.7	1.74	14.47
FEHLER 123	4773374o.8	148	322525.3		
INNERHALB VPN	3469678o9.8	646	5371o1.9		
GESAMT	7o85o7o71.9	683	1o36345.6		

S A C H R E G I S T E R

Abstimmung ("tuning") 197f

Adäquatheit, kognitive 262, 282f, 288f, 368f

Ähnlichkeit 63ff

Aktualgenese 93

Analyse
- globale 92ff
- lokale 92ff
- von Features 91, 92, 104

Antizipation 335 - 337, 346 - 347, 381 - 382

Bezugssysteme 378 - 379

Binnengliederung 199 - 203

Chunking 99ff, 115ff, 133f, 136ff, 156 - 157

Differenzierung 36

Diskrimination 36

echoischer Speicher 138

Erkennen (s. auch Identifikation)
- der Melodiekontur 47ff
- der zeitlichen Reihenfolge 30ff
- der zeitlichen Sukzession 31
- tonaler Muster 42ff, 140ff, 165ff, 204ff

Figur-Grund 60, 63 - 64, 65 - 68

Fokussierung 68, 110, 165f, 178f, 208

Fourieranalyse 197ff, 244ff

frame 205 - 206

Funktionalharmonik 275

Ganzheit, Ganzheitlichkeit 102 - 104

Gestalt 48, 64, 91

Gestalttheorie 81ff, 367 - 368

Grammatik, mentale 12

Handlungsplan 108

Handlungssteuerung 107

Heterarchie 111 - 112

Hicksches Gesetz 35

Hierarchie
- -modell 96, 97ff, 119ff, 145ff, 349ff
- propositionale 96
- Verarbeitungs- 96
-von Invarianzen 205 - 206

Hören, strukturelles 4 - 6, 387ff

Höreinstellungen, musikalische 1 - 2, 391ff

Identifikation, strukturelle 204f 246f

Informationstheorie der Musik 263ff

Informationsverarbeitung
- lineares Modell der 19, 20, 21, 26 - 27, 348ff
- auditive 22ff
- Stufen der 25 - 26, 37 - 41

Integration 53ff

Invarianz 9, 173, 192f, 194f, 204f, 244f, 378ff

Klickdetektion 333ff

Klicklokalisation 71 - 72, 122, 300ff

Kohärenz 56 - 58

Kompetenz 11

Konfiguration 47 - 52, 88ff, 104, 153, 165ff, 215f

Kontur 47ff, 182ff

Korrelationshypothese 10 - 11

Linguistik u. Musik 8, 261ff, 433ff

Maskierung 32, 138 - 141

Merkmale
- elementare 45 - 46
- konfigurative 23, 44 - 47, 88ff
- strukturelle 23

Merkmalsintegration 32 - 33

Modell, lineares 26 - 27

Musik
- Strukturbeschreibung der 6 - 9, 261ff, 433ff
- u. Sprache 6 - 9

Musikpsychologie 2, 18, 253ff, 367ff
- kognitive 253ff

Mustererkennung, auditive 22 - 24, 28ff, 42ff, 140ff

Nähe
- syntaktische 352 - 353, 357ff
- zeitliche 56, 352 - 353, 357ff

Objektpermanenz 68, 173

Organisation
- nach Ähnlichkeit 63 - 68
- rhythmische 58 - 62
- subjektive 34 - 36
- u. Kompatibilität 57

Organisationsalternativen 35

Ortsfrequenzanalyse 95 - 96

Performanz 11

Periode, musikalische 275 - 276

präattentiv 102

Präsenzzeit, psychische 54 - 55, 120f, 130ff

Psychologie, kognitive 25

Realität, psychologische 9 - 12

Redundanzausnutzung 135f

Regeln, musikalisch- syntaktische 269 - 270, 277ff, 370f

Repräsentation
- analoge 37 - 38, 190, 244
- ganzheitliche 98 - 99, 104
- hierarchische 73, 96-97, 97ff, 237
- in Form von "Strings" 37

Repräsentationsebenen 99ff

Resonanz 246ff

Rhythmus 58 - 62, 181, 383ff

Schema, kognitives 7, 129, 205, 335ff, 377 - 378

Segmentierung 53ff

Sprache u. Musik 6 - 9, 271ff

Stream Segregation 65 - 68

Struktur
- musikalische 8 - 12
- prosodische 58 - 62
- rhythmische 58 - 62

Strukturidentifikation 145ff, 152ff

Suffix-Effect 142 - 143

Syntax, musikalische 270ff, 283ff, 433ff

Transformation 49 - 51, 276 - 277, 285, 371f

Tuning 197f

Verarbeitungseinheiten 60, 68, 115ff, 153, 238

Verarbeitungshierarchie 96, 99ff

Wissen
- implizites 7, 259, 375 - 376
- musikalisches 72, 374ff

Zeitwahrnehmung 53

LEBENSLAUF

Geboren am 28.9.1948 in Bonn. Eltern: Prof. Dr. Hellmut Stoffer und Marlene Stoffer, geb. Kühle.

Besuch der Volksschule 1955 bis 1960 in Bonn und Köln; Besuch des Mathematisch-naturwissenschaftlichen Gymnasiums in Köln-Mülheim, dort Abitur Ostern 1968.

Studium der Psychologie und Pädagogik an der Ruhr-Universität Bochum von 1968 bis 1976; halbjähriges Stipendium der Heinrich-Hertz-Stiftung des Landes Nordrhein-Westfalen im Wintersemester 1971/72 für einen Forschungsaufenthalt in Philadelphia, U.S.A.; 1976 Diplom im Fach Psychologie.

1.4.1976 bis 30.4.1978 Tätigkeit als Wissenschaftliche Hilfskraft am Psychologischen Institut der Ruhr-Universität Bochum, Arbeitseinheit Kognitionspsychologie; seit dem 1.5.1978 als Wissenschaftlicher Assistent (m.d.V.b.).

Ich bin seit 1977 verheiratet.

Kleve, im März 1981

Thomas Stoffer

DER DRUCKER - Buscheystraße 138 - 4630 Bochum - Tel. (02 34) 70 41 16